# BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

## IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.**

**PCT** WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

2

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

| (51) International Patent Classification 6 : C12N 15/82, 15/29, 15/11, 5/10, A01H 5/00 | A1 | (11) International Publication Number: WO 98/28431 |
|---|---|---|
| | | (43) International Publication Date: 2 July 1998 (02.07.98) |

(21) International Application Number: PCT/GB97/03546

(22) International Filing Date: 24 December 1997 (24.12.97)

(30) Priority Data:
9626858.6 24 December 1996 (24.12.96) GB

(71) Applicant *(for all designated States except US)*: JOHN INNES CENTRE INNOVATIONS LIMITED [GB/GB]; Norwich Research Park, Colney Lane, Norwich NR4 7UH (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: JONES, Jonathan, Dallas, George [GB/GB]; 19 Waverley Road, Norwich, Norfolk NR4 6SG (GB). KLIMYUK, Victor Ivanovich [RU/GB]; 42 Grant Street, Norwich, Norfolk NR2 4HA (GB). DIRKS, Robert [BE/BE]; Kasteeldreef 23, B–3650 Dilsen–Stokkem (BE).

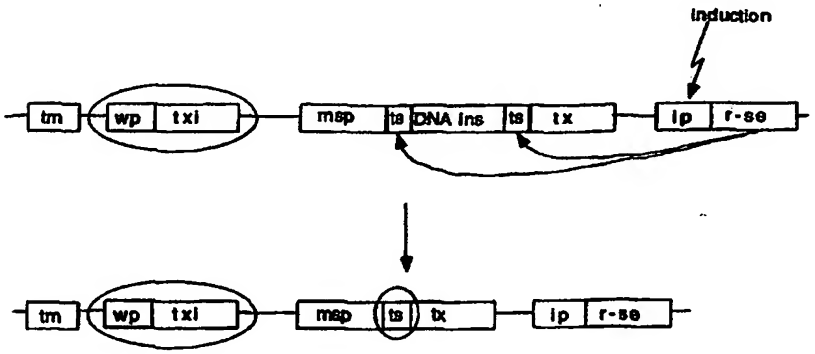
(74) Agents: WALTON, Seán et al.; Mewburn Ellis, York House, 23 Kingsway, London WC2B 6HP (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

**Published**
*With international search report.*
*Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.*



(54) Title: TRANSCRIPTIONAL REGULATION IN PLANTS




(57) Abstract

The present invention relates to regulatory sequences for meiosis–specific transcription of nucleic acid sequences or genes of interest and uses thereof. There is provided a promoter which confers meiosis–specific transcriptional regulation in plants. Such promoters may be used in improving transposon tagging efficiency, searching for apomictic mutants and/or the production of sterile plants. The present invention also provides a construct comprising a promoter of the present invention suitable for searching for apomictic mutants as well as for apomictic seed production.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| AL | Albania | ES | Spain | LS | Lesotho | SI | Slovenia |
| AM | Armenia | FI | Finland | LT | Lithuania | SK | Slovakia |
| AT | Austria | FR | France | LU | Luxembourg | SN | Senegal |
| AU | Australia | GA | Gabon | LV | Latvia | SZ | Swaziland |
| AZ | Azerbaijan | GB | United Kingdom | MC | Monaco | TD | Chad |
| BA | Bosnia and Herzegovina | GE | Georgia | MD | Republic of Moldova | TG | Togo |
| BB | Barbados | GH | Ghana | MG | Madagascar | TJ | Tajikistan |
| BE | Belgium | GN | Guinea | MK | The former Yugoslav Republic of Macedonia | TM | Turkmenistan |
| BF | Burkina Faso | GR | Greece | | | TR | Turkey |
| BG | Bulgaria | HU | Hungary | ML | Mali | TT | Trinidad and Tobago |
| BJ | Benin | IE | Ireland | MN | Mongolia | UA | Ukraine |
| BR | Brazil | IL | Israel | MR | Mauritania | UG | Uganda |
| BY | Belarus | IS | Iceland | MW | Malawi | US | United States of America |
| CA | Canada | IT | Italy | MX | Mexico | UZ | Uzbekistan |
| CF | Central African Republic | JP | Japan | NE | Niger | VN | Viet Nam |
| CG | Congo | KE | Kenya | NL | Netherlands | YU | Yugoslavia |
| CH | Switzerland | KG | Kyrgyzstan | NO | Norway | ZW | Zimbabwe |
| CI | Côte d'Ivoire | KP | Democratic People's Republic of Korea | NZ | New Zealand | | |
| CM | Cameroon | | | PL | Poland | | |
| CN | China | KR | Republic of Korea | PT | Portugal | | |
| CU | Cuba | KZ | Kazakstan | RO | Romania | | |
| CZ | Czech Republic | LC | Saint Lucia | RU | Russian Federation | | |
| DE | Germany | LI | Liechtenstein | SD | Sudan | | |
| DK | Denmark | LK | Sri Lanka | SE | Sweden | | |
| EE | Estonia | LR | Liberia | SG | Singapore | | |

# TRANSCRIPTIONAL REGULATION IN PLANTS

## Field of the invention

The present invention relates to meiosis in plants and more particularly the identification of regulatory sequences for meiosis-specific transcription of genes of interest and uses thereof.

## Background of the invention

Meiosis provides a mechanism by which a heterozygous individual can create large numbers of genotypically unique recombinant gametes. Chromosomes replicate during interphase, as in mitosis, and enter meiosis with two chromatids. During meiotic prophase I, chromosomes condense from the dispersed state typical of interphase, to form long thin threads in leptotene, and each acquires a proteinaceous axial core to which the two sister chromatids are attached. During zygotene, homologous chromosomes become aligned, forming the synaptonemal complex and, at pachytene, non-sister chromatids of the completely paired chromosomes recombine forming the chiasmata which become visible during diplotene. Two cell divisions follow - reductional and equational - resulting in four gametes, with each single chromosome as a potentially recombinant chromatid.

In yeast, molecular genetic analysis has revealed and led to the isolation of several genes that are essential for meiosis (Mitchell, 1995). For example, the DMC1 (Bishop et al., 1992) and RAD51 (Shinohara et al., 1992) genes are homologues of the E. coli recA gene, and appear to play a role not only in recombination-mediated homology-dependent pairing, but also in the strand exchange that results in chiasmata. Other genes, such as ZIP1, are required for synaptonemal complex formation (Sym et al., 1993 ).

In higher eukaryotes, molecular analysis of the mechanisms controlling chromosomal pairing has been

significantly more difficult than in yeast, since the systems are more complex and less easy to manipulate. However, several meiotic mutations have been identified in Drosophila (for review see Carpenter, 1993; 1994) and higher plants (Sears, 1976; Golubovskaya, 1989; Curtis and Doile, 1992; Golubovskaya et al., 1992, 1993; Tascheto and Pagliarini, 1993) and studied for their effect on meiosis. One such plant gene, Ph1, suppresses pairing of nonhomologous chromosomes in wheat and a Ph1-mutant background was used for the transfer and introgression of alien chromosome segments into wheat (Riley et al., 1968).

Plants have also provided an excellent cytological system for the study of meiosis (Gillies, 1984; Albini and Jones, 1987, 1988; Sherman and Stack, 1992; Albini, 1994; Schwarzacher and Heslop-Harrison, 1995), but there has been little investigation at the molecular level.

Lily anthers have offered the best system to study biochemical events that are correlated with the different stages of meiosis, because these stages are protracted and synchronised in adjacent flower buds, permitting the isolation of temporally regulated cDNA clones (Kobayashi et al., 1994). One such gene, LIM15, was expressed specifically in prophase I of meiotic cells and is extremely homologous to the yeast DMC1 gene (Kobayashi et al., 1993).

In both yeast and lily, DMC1 and RAD51 proteins colocalize during zygotene (Bishop, 1994; Terasawa et al., 1995). Early meiosis cDNA clones were also identified in wheat and maize by hybridisation to a Lilium meiosis-specific cDNA clone (Ji and Langridge, 1994).

The first genomic sequence of a recA-like plant gene, ArLIM15, with high degree of homology to that of LIM15, was recently described in Arabidopsis thaliana (Sato et al., 1995). However, no data showing the expression pattern of ArLIM15, as well as no characterization of meiosis-specific promoter have so far been reported by others. Sequence of upstream region of ArLIM15 gene, shown in Figure 4 of

Klimyuk and Jones 1996 (see below) contains predominantly the sequence of transposon-like element, Limpet1, (1874 bp) and only 260 bp of the promoter region, which is not sufficient to confer meiosis-specific expression of reporter gene.

For many applications it will be useful to drive transcription of different genes in specific parts of the plant at specific developmental stages. The promoters of plant meiotic genes are of extreme interest as they may provide transcriptional regulation of their genes during this very restricted developmental period. The isolation and characterization of these promoters enables the study and modification of fundamental processes taking place during early sporogenesis, as well as to study the impact of such modifications on more advanced stages of sexual reproduction in plants. Work by the present inventors outlining their objectives has been shown as a poster display (Klimyuk, V.I. et al. "The isolation and characterisation of the meiosis-specific Arabidopsis thaliana DMC1 gene". Abstracts of the 6th International Conference on Arabidopsis Research. June 7-11, 1995, Madison, Wisconsin, USA and Klimyuk, V.I. and Jones, J.D.G. "Identification of a transposon-like element, Limpet 1, in Arabidopsis thaliana". Abstracts of the 7th International Conference on Arabidopsis Research. June 23-27, 1996, Norwich, UK) and their work has been presented in a paper published after the priority date of the present application (Klimyuk, V.I. and Jones, J.D.G. "AtDMC1, the Arabidopsis homologue of the yeast DMC1 gene: characterization, transposon-induced allelic variation and meiosis-associated expression". Plant J (1997),11(1), 1-14) which is incorporated herein by reference.

## Summary of the invention

The primary aim of the inventors was to identify and isolate a plant promoter which confers meiosis-specific transcriptional regulation in plants.

They used the homologies between the lily LIM15 and the yeast DMC1 genes to design degenerate PCR (polymerase chain reaction) primers that amplified the Arabidopsis meiosis-specific DMC1 gene (designated AtDMC1). The AtDMC1 gene was completely sequenced and the transcript was characterised by RT(reverse transcriptase)-PCR. In situ hybridisation analysis showed that AtDMC1 expression is restricted to pollen mother cells in anthers and megaspore mother cells in ovules. A translational fusion was made between the AtDMC1 promoter and the GUS reporter gene. Transgenic Arabidopsis carrying the AtDMC1 promoter:GUS reporter gene fusion initiated GUS expression at the time of meiosis in both male and female lineages.

Sequence comparison of AtDMC1 and ArLIM15 (Sato et al., 1995) revealed that these genes, isolated from Landsberg erecta and Columbia ecotypes respectively, encode the same protein but are different in their promoter regions. We determined that the difference was caused by the insertion of a 1874 bp transposon-like element, designated Limpet1, into the promoter of ArLIM15. This finding revealed that previously published putative promoter region of ArLIM15 (Sato et al., 1995) predominantly contains Limpet1 sequences and would not be expected to be sufficient to confer meiosis-specific expression of the GUS reporter gene. The sequence of functional AtDMC1 promoter as well as the alignment of the promoter regions of the AtDMC1 and ArLIM15 genes upstream of their transcription start sites are shown in Figure 4A and B, respectively.

Reference herein to the AtDMC1 promoter being used in various ways, in accordance with the present invention, should be taken to be reference to a promoter including all or part of the promoter sequence shown in Figure 4(A), or a variant or derivative thereof, but excluding the promoter of AtLIM15 (Sato et al., 1995). A part (fragment), variant or derivative of the promoter sequence shown should be sufficient to confer meiosis-specific expression on a

heterologous sequence operatively linked, i.e. under the control of, the part, variant or derivative of the sequence shown. One or more fragments of the sequence may be included in a promoter according to the present invention, for instance one or more motifs may be coupled to a "minimal" promoter. Such motifs may confer meiosis-specific promoter function on a promoter which otherwise drives expression in a non-meiosis specific fashion.

According to a first aspect, the present invention provides a nucleic acid isolate comprising a promoter as indicated.

In a second aspect, the present invention provides a nucleic acid isolate comprising a promoter, the promoter comprising a sequence of nucleotides shown in Figure 4(A) and conferring meiosis specific expression on a sequence operably linked to the promoter. Restriction enzyme or nucleases may be used to digest the full-length nucleic acid shown, followed by an appropriate assay to determine the minimal sequence required for developmental specificity. A preferred embodiment of the present invention provides a nucleic acid isolate with the minimal nucleotide sequence shown in Figure 4(A) required for such specificity.

The promoter may comprise one or more sequence motifs or elements conferring developmental and/or tissue-specific regulatory control of expression. Other regulatory sequences may be included, for instance as identified by a mutation or digestion assay in an appropriate expression system or by sequence comparison with available information, e.g. using a computer to search on-line databases.

By "promoter" is meant a sequence of nucleotides from which transcription of DNA operably linked downstream (i.e. in the 3' direction on the sense strand of double-stranded DNA) may be initiated.

"Operably linked" means joined as part of the same nucleic acid molecule, suitably positioned and oriented for

transcription to be initiated from the promoter. DNA operably linked to a promoter is "under transcriptional initiation regulation" of the promoter.

The present invention extends to a promoter which has a nucleotide sequence which is allele, mutant, variant or derivative, by way of nucleotide addition, insertion, substitution or deletion of a promoter sequence as provided herein. Systematic or random mutagenesis of nucleic acid to make an alteration to the nucleotide sequence may be performed using any technique known to those skilled in the art. One or more alterations to a promoter sequence according to the present invention may increase or decrease promoter activity, or increase or decrease the magnitude of the effect of a substance able to modulate the promoter activity.

"Promoter activity" is used to refer to ability to initiate transcription. The level of promoter activity is quantifiable for instance by assessment of the amount of mRNA produced by transcription from the promoter or by assessment of the amount of protein product produced by translation of mRNA produced by transcription from the promoter. The amount of a specific mRNA present in an expression system may be determined for example using specific oligonucleotides which are able to hybridise with the mRNA and which are labelled or may be used in a specific amplification reaction such as the polymerase chain reaction. Use of a reporter gene facilitates determination of promoter activity by reference to protein production.

Therefore there is provided a nucleic acid molecule comprising

(a) the Arabidopis meiosis-specific DMC1 (AtDMC1) gene promoter; or

(b) a meiosis-specific promoter homologous to (a) but from another plant species; or

(c) a meiosis-specific promoter of the gene of a homologous DMC1 protein from another plant species; or

(d) a meiosis-specific promoter variant, mutant, allele or derivative of (a), (b) or (c); or

(e) a portion of (a), (b), (c) or (d) sufficient to confer meiosis-specific character to a promoter containing it.

In various embodiments of the present invention a promoter which has a sequence that is a fragment, mutant, allele, derivative or variant, by way of addition, insertion, deletion or substitution of one or more nucleotides, of the sequence of the promoter shown in Figure 4(A), has homology with the shown sequence which is at least about 5% greater than the homology that the ArLIM15 promoter sequence (Sato et al.) has with the sequence shown herein, preferably at least about 10% greater homology, more preferably at least about 20% homology, more preferably at least about 25% greater homology. The sequence in accordance with an embodiment of the invention may hybridise with the shown sequence but not the ArLIM15 promoter sequence under appropriately stringent selective hybridisation conditions. A promoter according to the invention may include one or more motifs that appear in Figure 4(A) and are able to confer meiosis-specificity on a promoter which contains them, and which are not present in the ArLIM15 promoter of Sato et al.

The present invention also includes meiosis-specific promoters that are homologous to the AtDMC1 gene promoter. Further, the present invention includes meiosis-specific promoters of the gene of a homologous DMC1 protein from another plant species. An homologous promoter or nucleic acid encoding an homologous DMC1 protein may show greater than 55% homology with the sequence of Fig. 4A or Fig. 5A or Fig. 5B, greater than 65% homology, greater than 75% homology, greater than 85% homology or greater than 95% homology. Such homology may be shown for a sequence of at least 20 nucleotide bases, at least 50 nucleotide bases, at least 100 nucleotide bases, at least 300 nucleotide bases or at least 500 nucleotide bases.

Further provided by the present invention is a nucleic acid construct comprising a promoter region or a fragment, mutant, allele, derivative or variant thereof as discussed able to promote transcription in a plant in a meiosis-specific manner, operably linked to a heterologous nucleic acid sequence, preferably a gene, e.g. a coding sequence. By "heterologous" is meant a gene other than the AtDMC1 coding sequence. Modified forms of AtDMC1 coding sequence may be excluded. Generally, the gene may be transcribed into mRNA which may be translated into a peptide or polypeptide product which may be detected and preferably quantitated following expression. A gene whose encoded product may be assayed following expression is termed a "reporter gene", i.e. a gene which "reports" on promoter activity.

The present invention also provides a nucleic acid vector comprising a promoter as disclosed herein. Such a vector may comprise a suitably positioned restriction site or other means for insertion into the vector of a sequence heterologous to the promoter to be operably linked thereto.

Suitable vectors can be chosen or constructed, containing appropriate regulatory sequences, including promoter sequences, terminator fragments, polyadenylation sequences, enhancer sequences, marker genes and other sequences as appropriate. For further details see, for example, *Molecular Cloning: a Laboratory Manual*: 2nd edition, Sambrook et al, 1989, Cold Spring Harbor Laboratory Press. Procedures for introducing DNA into cells depend on the host used, but are well known.

Thus, a further aspect of the present invention provides a host cell (which may be microbial or plant) containing a nucleic acid construct comprising a promoter element, as disclosed herein, operably linked to a heterologous nucleic acid sequence or gene. A still further aspect provides a method comprising introducing such a construct into a host cell. The introduction may employ any available technique well known to the person

skilled in the art.

The introduction may be followed by causing or allowing expression of the heterologous nucleic acid sequence or gene under the control of the promoter.

In one embodiment, the construct comprising promoter and nucleic acid sequence or gene integrated into the genome (e.g. chromosome) of the host cell. Integration may be promoted by including in the construct sequences which promote recombination with the genome, in accordance with standard techniques.

Many known techniques and protocols for manipulation of nucleic acid, for example in preparation of nucleic acid constructs, mutagenesis, sequencing, introduction of DNA into cells and gene expression, and analysis of proteins, are described in detail in *Short Protocols in Molecular Biology*, Second Edition, Ausubel et al. eds., John Wiley & Sons, 1992, the disclosure of which is incorporated herein by reference.

Nucleic acid molecules, constructs and vectors according to the present invention may be provided isolated and/or purified (i.e. from their natural environment), in substantially pure or homogeneous form, free or substantially free of a coding sequence, or free or substantially free of nucleic acid or genes of the species of interest or origin other than the promoter sequence. Nucleic acid according to the present invention may be wholly or partially synthetic. The term "isolate" encompasses all these possibilities.

An aspect of the present invention is the use of nucleic acid according to the invention in the production of a transgenic plant.

When introducing a chosen gene construct into a cell, certain considerations, well known to those skilled in the art must be taken into account. The nucleic acid to be inserted should be assembled within a construct which contains effective regulatory elements which will drive transcription. There must be available a method of

transporting the construct into the cell. Once the construct is within the cell membrane, integration into the endogenous chromosomal material either will or will not occur. Finally, as far as plants are concerned the target cell type must be such that cells can be regenerated into whole plants.

Plants transformed with the DNA segment containing the sequence may be produced by standard techniques which are already known for the genetic manipulation of plants. DNA can be transformed into plant cells using any suitable technology, such as a disarmed Ti-plasmid vector carried by Agrobacterium exploiting its natural gene transfer ability (EP-A-270355, EP-A-0116718, NAR 12(22) 8711 - 87215 1984), particle or microprojectile bombardment (US 5100792, EP-A-444882, EP-A-434616) microinjection (WO 92/09696, WO 94/00583, EP 331083, EP 175966, Green *et al.* (1987) *Plant Tissue and Cell Culture*, Academic Press), electroporation (EP 290395, WO 8706614) other forms of direct DNA uptake (DE 4005152, WO 9012096, US 4684611), liposome mediated DNA uptake (e.g. Freeman *et al. Plant Cell Physiol.* 29: 1353 (1984)), or the vortexing method (e.g. Kindle, *PNAS U.S.A.* 87: 1228 (1990d) Physical methods for the transformation of plant cells are reviewed in Oard, 1991, *Biotech. Adv.* 9: 1-11.

Agrobacterium transformation is widely used by those skilled in the art to transform dicotyledonous species. Recently, there has been substantial progress towards the routine production of stable, fertile transgenic plants in almost all economically relevant monocot plants (Toriyama, et al. (1988) *Bio/Technology* 6, 1072-1074; Zhang, *et al.* (1988) *Plant Cell Rep.* 7, 379-384; Zhang, *et al.* (1988) *Theor Appl Genet* 76, 835-840; Shimamoto, *et al.* (1989) *Nature* 338, 274-276; Datta, *et al.* (1990) *Bio/Technology* 8, 736-740; Christou, *et al.* (1991) *Bio/Technology* 9, 957-962; Peng, *et al.* (1991) International Rice Research Institute, Manila, Philippines 563-574; Cao, *et al.* (1992) *Plant Cell Rep.* 11, 585-591; Li, *et al.* (1993) *Plant Cell Rep.* 12,

250-255; Rathore, *et al.* (1993) *Plant Molecular Biology* 21, 871-884; Fromm, *et al.* (1990) *Bio/Technology* 8, 833-839; Gordon-Kamm, *et al.* (1990) *Plant Cell* 2, 603-618; D'Halluin, *et al.* (1992) *Plant Cell* 4, 1495-1505; Walters, *et al.* (1992) *Plant Molecular Biology* 18, 189-200; Koziel, *et al.* (1993) *Biotechnology* 11, 194-200; Vasil, I. K. (1994) *Plant Molecular Biology* 25, 925-937; Weeks, *et al.* (1993) *Plant Physiology* 102, 1077-1084; Somers, *et al.* (1992) *Bio/Technology* 10, 1589-1594; WO92/14828). In particular, *Agrobacterium* mediated transformation is now emerging also as an highly efficient alternative transformation method in monocots (Hiei *et al.* (1994) *The Plant Journal* 6, 271-282).

The generation of fertile transgenic plants has been achieved in the cereals rice, maize, wheat, oat, and barley (reviewed in Shimamoto, K. (1994) *Current Opinion in Biotechnology* 5, 158-162.; Vasil, *et al.* (1992) *Bio/Technology* 10, 667-674; Vain et al., 1995, *Biotechnology Advances*13 (4): 653-671; Vasil, 1996, *Nature Biotechnology* 14 page 702).

Microprojectile bombardment, electroporation and direct DNA uptake are preferred where Agrobacterium is inefficient or ineffective. Alternatively, a combination of different techniques may be employed to enhance the efficiency of the transformation process, eg bombardment with Agrobacterium coated microparticles (EP-A-486234) or microprojectile bombardment to induce wounding followed by co-cultivation with Agrobacterium (EP-A-486233).

Following transformation, a plant may be regenerated, e.g. from single cells, callus tissue or leaf discs, as is standard in the art. Almost any plant can be entirely regenerated from cells, tissues and organs of the plant. Available techniques are reviewd in Vasil et al., *Cell Culture and Somatic Cel Genetics of Plants, Vol I, II and III, Laboratory Procedures and Their Applications*, Academic Press, 1984, and Weissbach and Weissbach, *Methods for Plant Molecular Biology*, Academic Press, 1989.

The particular choice of a transformation technology will be determined by its efficiency to transform certain plant species as well as the experience and preference of the person practising the invention with a particular methodology of choice. It will be apparent to the skilled person that the particular choice of a transformation system to introduce nucleic acid into plant cells is not essential to or a limitation of the invention, nor is the choice of technique for plant regeneration.

Also according to the invention there is provided a plant cell having incorporated into its genome nucleic acid, particularly heterologous nucleic acid, as provided by the present invention. A further aspect of the present invention provides a method of making such a plant cell involving introduction of a vector including the sequence of nucleotides into a plant cell and causing or allowing recombination between the vector and the plant cell genome to introduce the sequence of nucleotides into the genome. The invention extends to plant cells containing nucleic acid according to the invention as a result of introduction of the nucleic acid into an ancestor cell.

Plants which include a plant cell according to the invention are also provided, along with any part or propagule thereof, seed, selfed or hybrid progeny and descendants. A plant according to the present invention may be one which does not breed true in one or more properties. Plant varieties may be excluded, particularly registrable plant varieties according to Plant Breeders' Rights. It is noted that a plant need not be considered a "plant variety" simply because it contains stably within its genome a transgene, introduced into a cell of the plant or an ancestor thereof.

In addition to a plant, the present invention provides any clone of such a plant, seed, selfed or hybrid progeny and descendants, and any part of any of these, such as cuttings or seeds. The invention provides any plant propagule, that is any part which may be used in

reproduction or propagation, sexual or asexual, including cuttings, seed and so on. Also encompassed by the invention is a plant which is a sexually or asexually propagated off-spring, clone or descendant of such a plant, or any part or propagule of said plant, off-spring, clone or descendant.

The AtDMC1 promoter may be used in meiosis-specific expression of a heterologous sequence, including a variety of cytotoxic genes, eg ribosome inactivating proteins (Lord et al., 1991) or cytotoxic RNase barnase (Goldman et al., 1994), DNA modifying enzymes including rare-cutting site-specific endonucleases, eg XbaI , I-SceI, HO endonuclease (Brenneman et al., 1996; Puchta et al., 1996; Chiurazzi et al., 1996; Haber, 1995), or recombinases, eg FLP recombinase (Kilby et al., 1995), recombinase from Zygosacharomyces rouxii (Onouchi et al., 1995), bacteriophage P1 cre recombinase (Osborne et al., 1995) as well as different transcription factors (Meshi & Iwabuchi, 1995; Ramachandran et al., 1994), protein kinases and phosphatases (Stone & Walker, 1995), cell cycle regulators (Ferreira et al., 1994; Dahl et al., 1995), which are normally not expressed during the time of meiosis. This may serve different purposes, such as ablation of meiotic cells and isolation of apomictic plants, designing an efficient homologous recombination system for plants, increasing meiotic recombination frequency, for introgression of alien chromosome segments into host plant, or altering normal events of cell cycle during the time of meiosis and producing male and female sterile plants. Easily assayed reporter genes, eg GUSA(Jefferson, 1987) or GFP (Sheen et al., 1995) under control of AtDMC1 promoter may be used as markers for detection of early meiotic events in plants, especially for the analysis of different mutations affecting meiosis.

The AtDMC1 promoter may be used for improving transposon tagging efficiency. The main factors determining the efficiency of transposon-based system in the tagging of

host genes are the frequencies of transposon excision and reinsertion, and the independence of transposition events. The last factor is crucial for a system with high level of excision and reinsertion events, as such system often yield clonal transpositions. Using the AtDMC1 promoter, driving the transpositions at the restricted stages of plant development, particularly at the early stages of meiosis, improves the efficiency of transposon tagging by producing only unique (independent) transposition events.

Therefore, the present invention provides a method of transposon tagging comprising the steps of creating construct comprising a promoter as described above operably linked to a transposase required for transposition of a transposable element; transforming a plant cell with said construct such that the transposase is driven by said promoter; and determining transposition events.

The AtDMC1 promoter may be used in searching for apomictic mutants or used by seed producers to produce seeds apomictically. Apomixis is the definition of asexual reproductive processes that occur in ovules of flowering plants (Koltunow, 1993). Mutants with fertilisation-independent seed development have been recently described for Arabidopsis (Ohad et al., 1996; Chaudhury et al.,1997). Apomictically produced plants are genetically identical with the female parental plants. Apomictic reproduction may therefore be beneficial for agriculture, as it may be an inexpensive way to preserve given genotype through successive generations.

Ablation of meiotic cells or changing their fate through altering events of meiotic cell cycle can be extremely beneficial for apomictic seed production, as it will eliminate sexually produced seeds from the progeny. Thus, the only seeds that will survive will be apomictic.

Schematic presentation of an exemplary construct suitable for searching for apomictic mutants as well as for the apomictic seeds production is shown in Figure 1A. The construct carries meiosis-specific promoter (msp) fused

through a DNA insertion sequence (DNA ins) to a modification nucleic acid sequence, preferably a gene (tx) encoding a cytotoxic protein. The DNA insertion sequence may be any DNA sequence (another gene, eg transformation marker or counter selectable marker) which does not confer transcription of cytotoxic gene and is sufficiently long to prevent transcription of cytotoxic gene from the meiosis-specific promoter. DNA insertion is preferably flanked with DNA sequences which are the target sequences (ts) for an activator, preferably a DNA recombinase (r-se), under control of inducible promoter (ip). The target sequences may be, for example, lox recombination sites with the same orientation in the case of the cre recombinase, or the 5' and 3' Ac ends for the Ac transposase.

The transcription of DNA recombinase gene may be controlled by an inducible promoter (ip), for example, promoter of heat shock inducible gene (Yoshida et al., 1995). Optionally, any other inducible system may be used, for example, copper-controllable (Mett et al., 1993), or glucocorticoid-controllable (Aoyama et al., 1995) gene expression system, but in this case the constructs design will be more complex, as such systems are transcription factor mediated.

Induction of the recombinase or transposase gene expression will lead to the elimination of the DNA insertion due to the homologous recombination between two flanking target sequences in the case of a cre/loxP- like system (one target sequence will be left behind) or complete excision of the DNA insertion together with both target sequences in case of Ac/Ds- like system. This will lead to the meiosis-specific transcription of cytotoxic gene and, consequently, meiotic cells death. For some powerful cytotoxins (barnase) the low expression level of inhibitor protein (txi) (barstar) would be an advantage, as it may prevent negative effect of possible leakage in cytotoxic gene expression.

This system may be useful, firstly for apomictic seed

production, as the induction of recombinase gene will lead to the ablation meiotic cells and viable seeds may be produced as the result of apomictic reproduction; and secondly, for the detection of apomictic mutants, especially in genotypes which are predisposed to apomixis when reproduction is not successful, e.g. citrus and grass species (Koltunow, 1993). In this case, seeds or plants or plant parts eg pollen carrying a construct as described in Figure 1A, may be mutagenised and the site- specific recombinase gene expression can be induced in self-progeny derived from mutagenised seeds. In this situation viable seeds will only be produced apomictically in an appropriate mutant background.

For increasing the efficiency of the system it may be an advantage to use easily screenable counter selectable marker, for example a bacterial cytochrome P450 (O'Keefe et al., 1994) or phosphonate monoester hydrolase (pehA) (Dotson et al., 1996) as the DNA insertion (DNA ins) (see Fig. 1A). These genes catalyse conversion of pro-herbicides into herbicides and any seedlings (potentially false positives) with inactive cytotoxic gene (tx) (because of the presence of DNA insertion) will be easily removed by the treatment appropriate pro-herbicide.

In this case the use of cre/lox -like system (i.e. one involving site-specific recombination) is preferable to Ac/Ds-like system, as the activation of first system will lead to the loss, not reintegration, of the DNA insertion.

Therefore, the present invention provides a method for producing seeds apomictically comprising the steps of modifying plant cells by incorporating into their genome a nucleic acid construct as described above so that the modification nucleic acid sequence is expressed in said meiotic cells thereby eliminating sexually produced seeds in a plant regenerated therefrom. The nucleic acid construct may be incorporated into the cell genome by breeding techniques well known to the person skilled in the art or by standard transformation techniques.

Also in accordance with the present invention there is provided a method of detecting apomictic mutants comprising the steps of generating a plant or plant part carrying a nucleic acid construct as described above; creating a variation for apomixis either in accordance with a breeding programme or in accordance with a mutagenesis programme; deriving a suitable self progeny population; inducing the activating gene in the self progeny as described above and detecting viable seeds produced apomictically. A suitable self progeny population will be understood by those skilled in the art to mean one where the phenotype of the mutation can be revealed e.g. where individuals are at least partially homozygous by selfing. Genetic variation could be created from breeding and selection of existing genetic stocks or by inducing such genetic variation, for example by mutagenesis.

In Figure 1B we present an exemplary construct permitting the removal of any unwanted DNA sequences from transgenic plants. It may be useful, for example, for eliminating genes conferring antibiotic resistance, which are required for selection of transformed plants, but are not necessary needed afterwards. Moreover, the presence of such genes can be damaging for the release of transgenic crops (or at least controversial, given the current political climate and food scares). The use of Cre/lox recombination system to remove selected genes from transgenic plants was already reported (Dale & Ow, 1991). The system described by Dale and Ow (1991) required the second round of transformation in order to introduce recombinase source under the control of constitutive promoter. In our construct meiosis specific promoter (msp) drives the transcription of site specific recombinase gene (r-se), which, e.g. as it was described above, eliminates DNA of interest (DNA ins) through the mechanism of homologous recombination during the first cycle of sexual reproduction in transgenic plant unlike the cre/lox-system. It will not require any additional rounds of transformation

as well as the use of inducible promoters to drive recombinase gene. The recombination event can be easily monitored in progeny of the primary transformant by placing DNA insertion between constitutive promoter (cp) and reporter gene (rg) or by using PCR screening.

Therefore, an aspect of the present invention is to prevent seed derived from sexual reproduction so that only apomictic seed are set (in plants where apomixis can occur). This would be of interest for species where making hybrid seed is uneconomic on a large scale. The hybrid seed may be made on a small scale and then be multiplied apomictically to produce large scale quantities of hybrid seed. Hybrid seed is desirable because hybrids generally perform better then inbreds.

Another aspect of the present invention is to identify plants with a tendency to form apomictic seed. This can be achieved in connection with a mutagenesis or breeding programme (see Fig. 1A). Further, the present invention also relates to a method for removing pieces of nucleic acid from transgenic plants (e.g. sequences just needed to transform/select transgenic cells/plants).

Aspects and embodiments of the present invention will now be illustrated, by way of example only, with reference to the accompanying figures. Further aspects and embodiments will be apparent to those skilled in the art. All documents mentioned herein are incorporated by reference.

## Brief description of the drawings

In the Figures:

**Figure 1** shows schematic representation of uses of meiosis-specific promoter in searching for apomictic mutants, maintaining the apomictic mechanism of seeds production and in removing of unwanted DNA sequences from transgenic plants. Abbreviations: tm - transformation marker; wp - weak promoter; tx- gene conferring toxic phenotype; txi - inhibitor of tx; msp - meiosis-specific

promoter; ts - target sequence for r-se gene product; DNA ins - any DNA sequence sufficiently large to block the meiosis-specific transcription of tx gene; cp - constitutive promoter; rg - reporter gene; ip - inducible promoter; r-se - gene conferring DNA recombination. Encircled parts of the constructs are optional.

Figure 2 shows the restriction map (A) of the AtDMC1 gene and T-DNA region of binary vector carrying pAtDMC1:GUS fusion. The positions of degenerate primers used to produce AtDMC1-specific probe A are shown by arrows. Probe B, which was used for RFLP mapping, corresponds to the 5.9 kb EcoRV fragment of genomic clone. The positions and sizes of 15 exons are shown by open boxes; the ATG and the TGA translation initiation and translation termination codons are also shown. The promoter region used for pAtDMC1:GUS translational fusion is shown as solid line. Figure 2(B) shows the schematic representation of the plasmid SLJ7753 T-DNA region introduced into A.thaliana in order to assess the specificity of the AtDMC1 promoter. The construct carries a pAtDMC1:GUS translational fusion and the NPT gene as a transformation marker.

Figure 3 shows nucleotide sequence of the AtDMC1 promoter and adjacent region. The nucleotide sequence from -450 to +330 is shown. The transcription start site is designated +1. The putative TATA box sequence is boxed. The alternative putative TATA box is overlined. Two nearly complete and two complete direct repeats are underlined. Consensus sequences of splicing are indicated by double underlining. The ATG codon in the second exon (shown in bold) was mutagenised in order to introduce an NcoI site for translational fusion with the GUS gene.

Figure 4 shows (A) DNA sequence of the AtDMC1 promoter including part of the transcribed region. Transcription start site located at the position 4691. (B) Alignment of the promoter regions of the AtDMC1 (top strand) and ArLIM15 (bottom strand) genes upstream of their transcription start sites. Only identical sequences in the alignment are shown.

Figure 5 shows DNA sequences of plant DMC1 homologues. Figure 5(A) shows partial genomic DNA sequence of barley DMC1 homologue (HvDMC1). Putative promoter region is located upstream of the translation start site (position 1599). Genomic DNA region amplified by the primers MEI1 and MEI4 is located between the positions 4001 - 4707.
Figure 5(B) shows partial genomic DNA sequence of tomato DMC1 homologue (LeDMC1). Genomic DNA region amplified by the primers MEI1 and MEI4 is located between the positions 466 - 1240.

Figure 6 shows a schematic presentation of the T-DNA regions of constructs SLJ11332 (A) and SLJ112315 (B).

Figure 7 shows alignment of the amino acid sequences of the AtDMC1, LIM15 and DMC1 proteins. Identical and conserved amino acid residues are boxed in black and grey, respectively. Gaps in the alignment are shown by dots. The positions and orientations of primers are shown by arrows. The vertical open arrowheads indicate the positions of introns in the AtDMC1 gene. The A and B motifs of the consensus sequence for the purine nucleotide binding site, detected by visual inspection, are underlined.

Figure 8 shows the restriction map of the promoter region of the ArLIM15 gene. The positions of the first five exons are shown by open boxes. The insertion of transposon-like element, Limpet1, flanked by two 9 bp direct repeats is also shown.

## Detailed description and exemplification of the invention

### Materials and Methods

*Plant material*

The plants used in this study were Arabidopsis thaliana Columbia and Landsberg erecta ecotypes. Plants were grown in the greenhouse at 25oC under 16 hrs of illumination and automatic watering conditions.

*Plant transformation*

Plant transformation of Arabidopsis thaliana (ecotype

Columbia) was performed as described (Bechtold et al., 1993). Seeds were harvested three weeks after the vacuum-infiltration, sterilised and screened for transformants on GM + 1% glucose medium (Valvekens et al., 1988) containing 50 mg l-1 kanamycin.

*DNA isolation and DNA gel blot analysis*

Genomic and cosmid DNA isolation from the plant tissue and E. coli respectively was performed as described (Klimyuk et al., 1995; Sambrook et al., 1989). The DNA was digested with restriction enzymes and DNA fragments were separated on 1% agarose gel, transferred to Hybond-N membranes (Sambrook et al., 1989), immobilised on the membranes by UV crosslinking (UV Stratagene stratalinker 2400) and subsequently baked on the membranes for 1 hour at 80oC. The hybridisation procedure was performed as described by Church and Gilbert (1984). DNA fragments for use as probes were gel-purified and were labelled using commercially available oligolabelling kit (Pharmacia).

*Subcloning and DNA Sequencing*

All subclonings and template preparations were done using the phagemid BlueScript (KS+) vector (Stratagene). The series of unidirectional 250-300 bp deletions were carried out for large inserts using the Erase-a-Base system (Promega). The sequencing reactions were performed by using DyeDeoxy Terminator cycle sequencing kit (Applied Biosystems). In some cases the PCR products were sequenced directly. Sequence analysis was carried out on ABI 373A DNA sequencer (Applied Biosystems). DNA contig, carrying AtDMC1 gene sequence, was built up by using Autoassembler programme (Applied Biosystems).

*Construction of pAtDMC1:GUS fusion*

The NcoI site was introduced by site-directed mutagenesis (Kunkel, 1985) at the position of the first ATG codon (shown in bold in Figure 3) in the second exon of 5.9

kb EcoRV subclone of the AtDMC1 gene to make plasmid SLJ7731. The sequence of the primer designated MUTA used for mutagenesis is: 5'- TAG AGC TGA AGALACTOSYL IGG AAC GAG CCC **CAT GGA** GCT CGT TGA GCG TGA-3'. The NcoI site is shown in bold. The final construct SLJ7731 was digested with NcoI and PstI restriction enzymes, gel purified and ligated with small NcoI-PstI fragment of SLJ4D4. The final plasmid SLJ7744, carrying pAtDMC1:GUS 3'ocs fusion in pBS(KS+) vector, was digested with HpaI and SmaI restriction enzymes. The large fragment, released by this digest, was gel purified and subcloned into HpaI site of binary vector SLJ491, based on pRK290. The final construct SLJ7753 (Figure 2B) was mobilised into Agrobacterium tumefaciens C58C1 strain harbouring the disarmed Ti plasmid pGV2260 (Deblaere et al., 1985) and used in transformation experiments.

*Histochemical GUS assay*

Transformed plants carrying T-DNA of SLJ7753 were XGluc stained at the different stages of development as previously described (Klimyuk et al., 1995). Squashes of X-Gluc stained, pigment-washed flower buds were prepared as described below. Specimens were placed in eppendorf tubes and vacuum-infiltrated with immersion oil in SpeedVac concentrator SVC100H (Savant). Then buds with remains of immersion oil were placed on slide and gently squashed with cover slip. The squashes were examined with a Zeiss Axiophot microscope.

The histochemical localisation of GUS activity in flower buds was performed on 10 mm cross-sections of Historesin-embedded plant material as it previously described (Dolan et al., 1994). For better visualisation of the tissue structure some of the cross-sections were stained for 1 min with 0.01% Safranin. The sections were examined with a Nikon Microphot-SA microscope equipped with a dark field condenser.

The following criteria were used to identify the

stages of flower development (Bowman, 1994): stage 9 - petal primordia stalked at base, pistil length (from the top of the stigma to the point of attachment to the receptacles) is 0.15 - 0.4 mm; stage 10 - stamen filaments begin to elongate, pistil length is 0.4 - 0.5 mm; stage 11 - stigmatic papillae appear, pistil length is 0.5 - 1.5 mm.

*RT-PCR and identification of cDNA ends (RACE)*

RT-PCR analysis of AtDMC1 expression was performed with MEI1U (5'-GGA GGG AAT GGA AAA GTG-3') primers using 1μg of total RNA from 12 day-old seedlings, leaves and floral buds as a template. The positions and orientations of the primers coincide with those of their degenerate homologues, MEI1 and MEI4 (Figure 2A). As an internal control, primers for Antirrhinum majus polyubiquitin mRNA (GenBank accession number X67957), kindly provided by Dr M. O'Dell, were used: 1392 (5'-CAG ATC TTT GTG AAG ACT CTG-3') And 1393 (5'-GGA CTC CTT CTG GAT GTT GTA-3'). Primer 1393, directed in antisense orientation was used for cDNA strand synthesis.

*RNA in situ hybridization*

Digoxigenin-labeling of RNA probes, tissue preparation and *in situ* hybridization were performed as described by Bradley et al. (1993) and Coen et al. (1990).

*X-ray treatment of Arabidopsis plants*

Twelve-day-old *Arabidopsis* plants grown in small (5cm in diameter) plastic petri dishes were exposed to 5 and 10 krad of ionizing irradiation. The ABB 6 MV linear accelerator served as a source of radiation. One hour after irradiation plants were used for X-Gluc staining and RNA isolation. Non-irradiated plants served as a control.

*RNA gel blot analysis*

RNA samples were separated on 1.4% agarose-formaldehyde gels as described by Ausebel et al. (1987). Agarose gel

was rinsed in several changes of sterile DEP-treated distilled water in order to remove formaldehyde and blotted overnight in 10xSSC to Hybond-N membrane. Membrane was carefully rinsed with deionized sterile water and RNA was immobilized to membrane and hybridized with probe as described above for DNA gel blot analysis.

**EXAMPLE 1**

*Isolation of the AtDMC1 gene*

In order to isolate the Arabidopsis thaliana DMC1 homologue, five different degenerate primers were designed corresponding to amino acid motifs conserved in LIM15 and DMC1 proteins: MEI1- 5′ GG(N) AA(GA) GT(N) GC(N) TA(CT) AT(ACT) GA 3′; MEI2 - 5′ GA(CT) AC(N) GA(GA) GG(N) AC(N) TT(CT) (CA)G(N) CC3′; MEI3 - 5′ A(GA)(TC) TT(TC) TG(TC) TG(N) C(GT)(TC) TC 3′; MEI4 - 5′ AC(N) GC(N) AC(GA) TT(GA) AA(CT) TC(CT) TC(N) GC 3′; MEI5 - 5′ GC(GA) TG(N) GC(N) A(GA)N AC(GA) TG(N) CC(N) CC 3′. The positions of primers and their orientations are shown in Figure 7. Different combinations of primers were used for PCRs with Arabidopsis genomic DNA as a template. The PCR reactions were performed with 0.05 μg of A. thaliana (Landsberg erecta) DNA as a template in a volume of 50 μl in the presence of 2 μM of each of the two selected primers in a buffer containing 250 μM dNTPs (Pharmacia) 10 mM Tris-HCl, pH8.3, 50 mM KCl, 2.5 mM MgCl2, 0.05% Nonidet P-40, and 2.5 units of "AmpliTaq" thermostable DNA polymerase (Perkin Elmer Cetus). Cycling conditions were: 94 oC for 15 sec; 50 oC for 30 sec; 72 oC for 2 min; 35 cycles, followed by a 10 min extension at 72 oC. The largest PCR products were reamplified with nested primers. Only two sets of primers, MEI1 - MEI5 and MEI1 - MEI4, proved effective. Primers MEI2 and MEI3 did not amplify the expected size class either together or in combination with other primers. Subsequent analysis showed that they anneal to parts of the cDNA sequence that are interrupted by introns in genomic DNA (Figure 7). The major 679 bp PCR band, obtained after the reamplification with

MEI1 and MEI4 primers of the first round PCR product (MEI1 and MEI5 primers), was blunt-ended by treatment with T4 polymerase, subcloned into the EcoRV site of pBS(KS+) and sequenced. The fragment, designated as probe A (Figure 2A), encoded 111 amino acid sequence with 93% of identity to the homologous part of LIM15 15 protein (Kobayashi et al., 1993).

Two unique primers, MEI1U (5' GGA GGG AAT GGA AAA GTG 3') and MEI4U (5' GCA ACG TTG AAC TCC TCT GCA AT 3'), that annealed to the same location as their degenerate homologues (Figure 2A), were synthesised and used for the PCR screening of DNA pools prepared from 68 plates (384 clones per plate) of an four genomic equivalents Arabidopsis cosmid library. The library was made on the basis of CLD04541 binary vector and was kindly provided by C. Lister and C. Dean (Cambridge Laboratory, JIC, Norwich). The probing of filter replicas from positive plates with probe A recovered five cosmid clones. One of the clones, called 64/23/C, was used for the restriction mapping and subcloning of 5.9 kb EcoRV and 3.6 kb ClaI overlapping fragments into the pBS (KS+) vector. Both strands of two overlapping fragments, encompassing 8 kb of cosmid insert, were completely sequenced.

The DNA sequencing and sequence analysis was performed as described above. Database searches with the BlastX program (Altschul et al., 1990) revealed a gene within this region, designated AtDMC1, whose highest homology was to the ArLIM15 protein followed by LIM15 and its human and mouse homologues as well as several RAD51 homologues and other RecA-like proteins, including yeast DMC1, from different eukaryotic organisms. The GenBank accession number for the AtDMC1 gene is U76670. DNA gel blot analysis of AtDMC1 revealed that this is a single copy gene.

*Molecular mapping*

Labelled 5.9 kb EcoRV fragment of AtDMC1 genomic clone

(designated as probe B; Figure 2A) was initially used to probe Southern blot of Arabidopsis thaliana (ecotypes Landsberg erecta and Columbia) DNA to detect specific RFLPs. AtDMC1 was shown to be polymorphic with EcoRI, EcoRV, XbaI, HindIII, HpaI and BstEI enzymes between A. thaliana Landsberg erecta and Columbia ecotypes. Subsequently, Southern blots of EcoRI digested DNAs isolated from 41 recombinant inbred (RI) lines between the ecotypes Landsberg erecta and Columbia (Lister and Dean, 1993) were hybridised with probe B. The molecular mapping of AtDMC1 gene was carried out by using the program MAPMAKER V.1.0 (Lander et al., 1987), and data for the segregation of 92 single-copy sequences covering the five Arabidopsis chromosomes.

The RFLP for the AtDMC1 was mapped to the top arm of chromosome 3 between the m560B2 and g4711 molecular markers at the distances of 4.8 cM and 7.7 cM from each of the markers respectively. In addition mapping of AtDMC1 using recombinant inbred lines of Landsberg erecta and Columbia ecotypes led to the identification of a single locus, and the conclusion that AtDMC1 and ArLIM15 are allelic.

*Characterisation of the AtDMC1 transcript by RT-PCR*

Screening of an Arabidopsis inflorescence cDNA library made using the ZAPII vector (distributed by J. Dangl) with probe A did not reveal any positive clones, probably due to the low abundance or transitory presence of the AtDMC1 transcripts. Comparison of conceptual translations of the AtDMC1 gene to the LIM15 cDNA sequence showed a very high degree of amino acid sequence homology (more than 80% identity). On the basis of this homology 14 exons and 13 introns were predicted within the AtDMC1 gene. In order to confirm the prediction and to define precisely the boundaries of exons, RT-PCR , 5′ RACE and 3′ RACE were carried out as described (Frohman et al., 1988; Frohman, 1989) using total RNA from floral buds as a template.

Total RNA was prepared from Arabidopsis 0.1 - 1 mm

floral buds and young leaves by a standard method (Harpster et al., 1988) or by method described by Napoli et al., (1990). Both methods produced good quality RNA, but the last one was preferred when small quantities of total RNA were required.

For RT-PCR complementary DNA synthesis was carried out with 2 μg of total RNA and primer MEI6U (5' ATC CTT CGC GTC AGC AAT GCC 3'). Second strand synthesis and PCR were carried out with primers 5RN (5' ATG CAG CTC GTT GAG CGT GAA 3') and MEI6U. PCR mixture was heated for 5 min at 95oC, followed by 35 cycles of amplifications (94°C, 40 sec; 56°C, 1 min; 72°C, 2 min) and 10 min final extension at 72°C.

According to the prediction, the primers should amplify 996 bp and 2586 bp fragments from the cDNA and genomic DNA respectively. Two expected bands were observed after the electrophoretic separation of RT-PCR products and the 996 bp fragment was cloned and sequenced. The sequence confirmed the prediction, except that minor corrections were necessary to the predicted boundaries of some exons.

Further steps were undertaken to identify the full length transcript, in particular the transcription start site. The GENEFINDER program (developed by S. Klostermann, Max-Planck Institute, Martinsried) predicted an additional short exon 5' to the first of the 14 exons that were initially identified.

For identification of the transcription start site, 5' RACE complementary DNA synthesis was performed with 5 μg of total RNA and primer 5R2 (5' TCA GCA GCT TCA CAG ATT TTG 3'). Second strand cDNA synthesis and first PCR amplification were done with QI, QT (Frohman, 1989) and 5R2N (5' TCA ACT TTG GCC TCA GAT AAA C 3') primers. Reamplification was done with QI and 5R2N1 (5' TTC TTG GTA TGC ATC ATG AGALACTOSYL IGG 3') primers. Several 5' RACE products were size-selected by gel electrophoresis, subcloned and sequenced as described above for genomic DNA. The beginning of the longest was presumed to correspond to

the transcriptional start site, designated +1 in Figure 3. This result confirmed the existence of additional exon and allowed us to define the promoter of the AtDMC1 gene.

The promoter contains some interesting motifs upstream of the putative TATA box. Four direct repeats were found at the positions -285 to -397 (Figure 3). Two nearly complete repeats, 9 bp and 11 bp, are flanked by two complete 15 bp repeats. Interestingly, three of the repeats contain the short palindromic repeat ATGCAT at their 3' ends. Transfac v 3.2 database search for homology with known transcription factor binding sites using TESS - Transcription Element Search Software (http/ /agave. humgen. upenn. edu/ tess/ index. html) revealed that these repeats contain putative transcription factor binding domains. The repeats contain 6 - 11 bp long sequences that show homology to the transcription factor binding sites for quail transcription factor EFII (Sealey & Chalkley. 1987), human glucocorticoid receptor (Haerd et al., 1990) and Xenopus octamer-binding factor (Tebb & Mattaj, 1989).

For 3' RACE complementary DNA synthesis was carried out with 5 $\mu$g of total RNA and QT primer (Frohman, 1989). Second-strand cDNA synthesis and first PCR were done with QO and GSP1 (5' TCT GGG AAA ACC CAA TAA 3') primers. Reamplification was done with QI and GSP2 (5' GCA CAT ACC CTT TGT GTC 3') primers. Results of 3' RACE revealed a 260 bp untranslated region (excluding the polyA tail). The full length mRNA transcript sequence was inferred from compiling the RT-PCR and RACE data.

This transcript codes 344 amino acids putative AtDMC1 protein which exhibits significant sequence similarity to lily LIM15 (Kobayashi et al., 1993) and yeast DMC1 (Bishop et al., 1992) meiotic proteins. The optimal alignment showed 84.3% of amino acid identity and 93.6% amino acid similarity between AtDMC1 and LIM15 and 51.8% of amino acid identity and 71.1% similarity between AtDMC1 and DMC1. AtDMC1 exhibits a somewhat lower level of homology with the yeast RAD51, a protein required for mitotic and meiotic

recombination and resistance to ionising radiation, (48.5% and 70.6% of identity and similarity respectively). AtDMC1, like other RecA-like DNA strand-exchange proteins (Kowalczykowski and Eggleston, 1994), possesses consensus ATP binding sites (Walker et al., 1982).

Comparison of the AtDMC1 and the ArLIM15 (Sato et al., 1995) transcripts and genomic sequences confirmed positions for most of the exon/intron junctions, but the borders of intron 14 of ArLIM15 gene appears to have been determined incorrectly. As a result, two amino acids, alanine (A) and glutamic acid (E) (positions 326 and 327 for AtDMC1 protein) were excluded from the predicted protein sequence of ArLIM15 (Sato et al., 1995). It is unlikely that splicing for this particular intron is different between Columbia and Landsberg erecta ecotypes. There is amino acid substitution at the position 103 (leucine in Columbia, serine in Landsberg erecta ). There is also a difference in the sizes of transcripts: AtDMC1 transcription start site is 4 bp upstream of that for the ArLIM15 and the last exon is 75 bp longer.

*In situ hybridization analysis of AtDMC1 expression*

In order to test whether or not the AtDMC1 gene is expressed at the time of meiosis, in situ hybridisation analysis using cross-sections of the whole inflorescence and DIG-labelled antisense AtDMC1 RNA as a probe was carried out. Digoxigenin-labelling of RNA probes, tissue preparation and in situ hybridisation were performed as described by Bradley et al. (1993) and Coen et al. (1990).

The expression of AtDMC1 gene in whorl 3 is restricted to pollen mother cells. No expression was detected in tapetum. In whorl 4 hybridisation signal was restricted to megaspore mother cells of ovules. DIG-labelled sense AtDMC1 RNA was used as a negative control and did not reveal any hybridisation signal at the stages of flower development at which the expression of AtDMC1 might take place. Weak nonspecific signal was detected in the mature pollen grains

following hybridization with sense and antisense AtDMC1 RNAs (data not shown). No signals were detected in postmeiotic ovules and developing embryos as well as in the analysed vegetative tissues.

*Characterization of expression of a pAtDMC1:GUS fusion*

The meiosis-associated expression of the AtDMC1 led us to investigate whether the AtDMC1 promoter could direct meiosis-associated expression of the GUS reporter gene. A translational fusion was made between the putative AtDMC1 promoter and coding sequences of the GUS reporter gene (see: *Construction of pAtDMC1:GUS fusion* in *Materials and Methods*).

The fusion consists of a 3.3 kb DNA fragment containing the AtDMC1 promoter fused in frame with GUS at the position of the methionine residue located in the second exon (Figures 2b and 3). The sequence of the 3.3 kb DNA fragment of the AtDMC1 gene used to drive meiosis-specific expression of a GUS reporter gene and its alignment with previously published sequence of the ArLIM15 gene are shown in Figure 3. As a result the GUS protein carries 13 AtDMC1 amino acid residues at its amino terminus. These residues appear to be neutral with respect to GUS activity. A schematic representation of the T-DNA region carrying the pAtDMC1:GUS fusion is shown in Figure 2b.

Eight primary transformants with this pAtDMC1:GUS fusion were obtained and analysed for the presence of GUS expression patterns. Three primary transformants did not reveal any GUS activity, one showed ubiquitous GUS expression and four transformants revealed GUS expression patterns which were restricted to the whorl 3 and whorl 4 of flowers. No GUS expression was detected in roots, leaves and stems of these four transformants, except that one of them exhibited GUS expression in damaged tissues.

There was also X-Gluc staining in the receptacles of some of the open flowers, but this pattern is very common

for plants carrying the GUS gene and may be considered as non-specific (Klimyuk et al., 1995).

The GUS expression initially appears in the anthers of approximately 0.2 mm long inflorescence buds and later in the carpels of more advanced, approximately 1 mm long, buds. Meiosis in anthers and carpels does not coincide (Bowman, 1994). The beginning of meiotic prophase I in anthers takes place at stage 9 of flower development, while the meiotic events in carpels do not occur until stage 11 (Bowman, 1994). Cross-sections of X-Gluc-stained buds at the different stages of development showed that GUS expression first appeared in pollen mother cells of anthers from inflorescence buds early in stage 9. In late stage 9 GUS staining increased dramatically. GUS expression in ovaries was first detected at stage 11 of flower development.

This temporal and spatial coincidence of reporter gene expression driven by the AtDMC1 promoter with the stages of floral bud development corresponding to the time of meiosis provides indication that AtDMC1 promoter can drive meiosis-associated GUS expression. The presence of residual GUS activity adjacent to the main sites of localisation in anthers and ovaries is the result of artefactual indigo blue dye formation in surrounding tissue. The specificity of dye localisation can be improved by using 0.2 - 1 mM potassium ferrocyanide/ferricyanide in staining solution (Jefferson, 1987). However, this we found compromised the sensitivity of the protocol resulting in loss of the weak GUS expression patterns at the early stages of meiosis. Inflorescences of nontransformed Arabidopsis plants were used as a negative control and they did not reveal any patterns of X-Gluc staining.

*AtDMC1 and ArLIM15 genes are different within the promoter regions*

Alignment of the AtDMC1 and ArLIM15 (Sato et al., 1995) gene sequences revealed that these genes, isolated

from ecotypes Landsberg erecta and Columbia respectively, are virtually identical throughout their respective coding regions, but there are significant differences in the 5' sequences. The sequences diverge 219 bp upstream of the transcriptional start site of AtDMC1 gene.

Different combinations of primers ATM2 (5' GCA ACT GAA TTT GTT TTC GTT TG 3'), ATM1 (5' TTG ATT AGT GGA TCC GCA AAC AA 3') and AR2 (5' TAG ATG AAA CGA GTT TGA CAC ATG 3') were used for PCR amplification of genomic DNA, isolated from Landsberg erecta and Columbia ecotypes of A. thaliana. The PCR amplification was performed as it described above for isolation of genomic clones, except that each primer concentration was 0.1 $\mu$M and cycling conditions were: 94°C for 20 sec; 58°C for 20 sec; 72°C for 2 min; 35 cycles, followed by a 10 min extension at 72°C. Primers T1 (5' GGG AAT GTT CCA ATA TAA G 3') and T2 (5' GAG AAT ATT ACA CTC TTA AA 3') were used for amplification of Limpet1 sequences. The set of primers AR2 - ATM2 produced the expected 446 bp PCR band for Columbia ecotype, but nothing for Landsberg erecta, whereas primers ATM1 - ATM2 produced the expected 342 bp PCR product for Landsberg erecta and 2.2 kb PCR product from Columbia ecotype. The positions and orientations of primers are shown in Figure 8.

This result confirmed that the AtDMC1 and ArLIM15 sequences described above are not derived from chimeric clones and reflect genuine differences in gene structure. To determine the nature of the rearrangement, sequence of the ArLIM15 promoter region was extended. This indicated that the difference between the AtDMC1 and ArLIM15 promoter regions was caused by a 1874 bp DNA fragment present in the ArLIM15 gene but absent in the AtDMC1 (Figure 8). The fragment is flanked by two 9 bp direct repeats and has 26 bp imperfect inverted repeats, with 73% of identity, at its termini; internal to these are two shorter inverted repeats (GenBank accession number U76697). The general structure of the fragment suggests that it is a transposable element and that it is probably a member of the class of transposons

that exhibit DNA-mediated transposition. The two 9 bp direct repeats at the ends of the element probably represent a target site duplication. Consistent with this suggestion such a duplication is absent within the AtDMC1 promoter.

The element, which we have designated Limpet1 (LIM promoter entrenched transposon-like element), was used as a probe to blots of genomic DNA from three different ecotypes of Arabidopsis thaliana. Primers T1 and T2 were used to produce Limpet1 hybridisation probe by reamplification of the ATM1 - ATM2 PCR product from ecotype Columbia. DNA gel blot analysis indicated that Columbia ecotype contains Limpet1 and at least two additional closely related elements, but no Limpet1 or related sequences were found in Landsberg erecta .

In order to assess whether or not the insertion of Limpet1 affects the expression pattern of ArLIM15, relative to the pattern exhibited by AtDMC1, RT-PCR was performed with MEI1U and MEI4U primers, using equal amount of total RNA isolated from floral buds and leaves of both ecotypes. The results of DNA gel blot analysis of RT-PCR products suggest that two ecotypes exhibit very similar and relatively high levels of expression in inflorescence tissue and similar but much lower levels of expression in the leaves.

*Isolation of barley and tomato DMC1 homologues*

Degenerate primers MEI1, MEI4 and MEI5, which proved to be efficient for isolation of AtDMC1 gene, were used to generate gene-specific probes for barley and tomato DMC1 homologues designated as HvDMC1 and LeDMC1, respectively. The first amplification was performed with MEI1 - MEI5 primers at the conditions described above for the isolation of the AtDMC1 gene-specific probe. The first PCR product was 10 times diluted with sterile distilled water and 1$\mu$l of it was used as a template for re-amplification with MEI1 - MEI4 set of primers. Agarose gel electrophoresis of PCR

products showed single major DNA band of approximately 700 and 800 bp long for barley and tomato, respectively. Gel-purified DNA bands were blunt-ended, cloned into the EcoRV site of pBS(KS+) and sequenced. Comparative analysis of the clones revealed that they encode amino acid sequences with more than 90% of identity to the homologous parts of LIM15 and AtDMC1 proteins. These clones encoding part of the LeDMC1 and HvDMC1 genes, were used to screen the tomato cosmid (Dixon et al., 1996) and barley lambda (Stratagene) genomic libraries, respectively. Partial sequences of genomic clones recovered from the screens are shown in Figure 5. Database similarity search using BlastX program (Altshul et al., 1990) helped to identify the translation start codon of HvDMC1 gene. Only HvDMC1 genomic clone was sequenced far enough upstream of the coding region to confer at least 1.5 kb of the putative promoter sequences (Figure 5a).

Success in isolation of DMC1 homologues from Arabidopsis, barley and tomato using set of MEI1, MEI4 and MEI5 degenerate primers is a convincing example that these primers can be used to isolate DMC1 homologues from monocot and dicot plant species. Comparative analysis of the coding regions of plant DMC1 homologues including soybean DMC1 homologue cDNA ( Database accession number U66836, direct submission 13 August 1996) and recently isolated partial sequence of rice LIM15-like gene (Database accession number U85613, direct submission 16 January 1997) demonstrated that they are at least 70% identical. The most variable part of the coding sequences is located at the 5' end of cDNA and encompass approximately 80 bp from the translation start site. It also revealed that plant DMC1 homologues have conserved exon/intron structure, which explains the success of using MEI1, MEI2 and MEI4 primers in different species.

**EXAMPLE II**

*Construction of pAtDMC1:barnase fusion for meiotic cells*

*ablation.*

The plasmid SLJ7744, carrying pAtDMC1:GUS 3'ocs fusion in pBS(KS+) vector, was digested with NcoI and EcoRV, large fragment was gel-purified and ligated with NcoI - EcoRV fragment of pJB142 carrying barnase-barstar-CaMV polyA fusion. The final construct was XbaI digested and 6 kb gel-purified fragment, carrying pAtDMC1:barnase-barstar-CaMV polyA signal, was cloned into the XbaI site of the binary vector SLJ755I5. The final construct SLJ11332 (Figure 6a) was mobilised into Agrobacterium tumefaciens C58C1 strain and used in transformation experiments.

*Construction of vector carrying AtDMC1 cDNA fused to potato virus X (PVX) amplicon*

The first 450 bp of the AtDMC1 cDNA were amplified with Cla1/1 (5'- CAAAATTCTATCGATCTCACTCTTCCAAGCTTA-3') and Cla1/2 (5'- CAAAAGCCTCTGTGATCGATGAGGTTTCAATTCCACC - 3') primers with introduced Cla1 sites (shown in bold). PCR fragment was digested with Cla1, gel-purified and cloned into the Cla1 site of binary vector pVDH401 carrying PVX amplicon (Angell & Baulcombe, 1997). The final construct SLJ112315 (Figure 6b) was mobilised into Agrobacterium strain and used in transformation experiments as it described in *Materials and Methods.*

It was shown that amplicon-mediated gene silencing can be used as an efficient tool to suppress endogenous RNA sharing homology with the transgene (Angell & Baulcombe, 1997). Considering that yeast dmc1 mutant fails to form normal synaptonemal complex and, as a result, produces a very low percentage of viable spores (Bishop et al., 1992), amplicon-mediated silencing of the AtDMC1 expression can be an alternative way to switch off the mechanism of sexual reproduction in Arabidopsis. Coding sequences of Arabidopsis, tomato, barley and rice (Database accession number U85613, direct submission 16 January 1997) homologues share 70 - 80% of homology.

Considering this, it is possible that amplicon- based AtDMC1 cDNA can efficiently silence DMC1 homologues expression in other plant species. Alternatively, host DMC1 homologue cDNA can be used in amplicon to switch off the mechanism of sexual reproduction and to achieve apomictic seed production.

The remarkable specificity of the AtDMC1 which confers tight developmental regulation of reporter gene expression in whorl 3 and whorl 4, may serve as a model system to study the mechanism of such regulation. Four direct repeats identified upstream of AtDMC1 TATA box (Figure 3) may play an important role in developmental regulation of AtDMC1 gene expression. Functional dissection of the promoter and site-directed mutagenesis may help to identify cis-regulatory sequences controlling the transcription of the AtDMC1 gene and to clarify possible involvement of direct repeats in such regulation. Such sequences may be used as "baits" in one-hybrid system.

Identification of the minimal region of the AtDMC1 promoter possessing all the sequences necessary to drive meiosis-specific transcription can be achieved by preparing constructs carrying truncated AtDMC1 promoter fused to the GUS reporter gene, transforming them into Arabidopsis and testing them for the ability to confer meiosis-specific GUS expression. These experiments will produce a truncated meiosis-specific promoter or even meiosis-specific enhancer sequences which will be more suitable for making constructs shorter than its present 3 kb long version. Meiosis-specific enhancers may be fused to any heterologous "minimal" promoter, for example, with -67 bp 35S promoter, which is not able to drive transcription, but contains RNA polymerase binding site.

The present inventors have made constructs carrying different types of fusion of the AtDMC1 promoter to different transposase genes (Ac, Spm) as well as to the

Tnt1 tobacco retrotransposon.

In the case of Tnt1 retrotransposon, part of 5'LTR upstream of retrotransposon TATA box was replaced with the AtDMC1 promoter sequences located upstream of the AtDMC1 TATA box. Transcriptional fusion within nontranslated leader of AtDMC1 gene was made for Spm transposase and pAtDMC1:10ATG.Ac transposase translational fusion was performed within the second exon of AtDMC1 gene, as it described for pAtDMC1:GUS construct (Figure 2B).

Isolation of the promoters of DMC1 homologues from other plant species and testing their abilities as well as the ability of pAtDMC1 to drive meiosis-specific transcription in heterologous plant systems can easily be envisaged.

References

Albini SUBJECT-MATTER (1994). Plant J. 5: 665-672.

Albini SUBJECT-MATTER & Jones GH (1987). Chromosoma 95: 324-338.

Albini SUBJECT-MATTER & Jones GH (1988). Genome 30: 399-410.

Altschul SF, et al (1990). J. Mol. Biol. 215:403-410.

Angell SUBJECT-MATTER & Baulcombe DC (1997). EMBO j. 16:3675-3684.

Aoyama T, et al. (1995). Plant Cell 7:1773-1785.

Bishop DK (1994). Cell 79:1081-1092.

Bishop DK, et al. (1992). Cell 69:439-456.

Bowman J (1994) Arabidopsis. An atlas of morphology and development. (New York: Springer-Verlag New York, Inc.), pp. 133-332.

Brenneman M, et al. (1996). Proc. Natl. Acad. Sci. USA 93:3608-3612.

Carpenter ATC (1993). BioEssays 16:69-74.

Carpenter ATC (1994). Cell 77: 959-972.

Chaudhury AM et al. (1997), Proc. Natl. Acad. Sci. USA 94:4223-4228.

Chiurazzi M, ET AL. (1996).
Church GM & Gilbert W (1984). Proc. Natl. Acad. Sci. USA 81:1991-1995.
Curtis CA & Doyle GG (1992). J. Heredity 83:335-341.
Dahl M, et al. (1995). Plant Cell 7:1847-1857.
Dale EC & Ow DW (1991). Proc. Natl. Acad. Sci. USA 88:10558-10562.
Deblaere R, et al. (1985). Nucl. Acids Res 13:4777-4788.
Dixon MS, et al.(1996). Cell 84:451-459.
Dolan L, et al. (1994). Development 120:2465-2476.
Dotson SB et al., (1996). Plant J. 10:383-392.
Ferreira P, et al. (1994). Proc. Natl. Acad. Sci. USA 91:11313-11317.
Frohman MA (1989). Methods Enzymol. 218:340-356.
Frohman MA, et al. (1988). Proc. Natl. Acad. Sci. USA 85:8998-9002.
Gillies CB (1984). CRC Rev. Plant Sci. 2:81-116.
Goldman MH, Goldberg RB & Mariani C (1994). EMBO J 13:2976-2984.
Golubovskaya IN (1989). Adv. Genet. 26:149-192.
Golubovskaya I, et al. (1992). Dev. Genet. 13:411-424.
Golubovskaya I, et al. (1993). Genetics 135: 1151-1166.
Haber JE (1995). BioEssays 17:609-619.
Haerd T, et al.. (1990). Science, 249:157-160.
Harpster MH, et al. (1988). Mol. Gen. Genet. 212:182-190.
Jefferson RA (1987) . Plant Mol. Biol. Rep. 5:387-405.
Ji L-H & Langridge P (1994). Mol. Gen. Genet. 243: 17-23.
Jucovic M, & Hartley RW (1996). Proc. Natl. Acad. Sci. USA 93:2434-2437.
Kilby NJ, et al. (1995). Plant J. 8:637-652.
Klimyuk VI, et al. (1995). Mol. Gen. Genet. 249:357-365.
Klimyuk VI & Jones JDG (1997). Plant J. 11:1-14.
Kobayashi T, et al. (1993). Mol. Gen. Genet. 237: 225-232.
Kobayashi T, et al. (1994). DNA Res. 1:15-26.
Koes R, et al. (1994). Annu. Rev. Biochem 63:991-1034.
Koltunow AM (1993). Plant Cell 5:1425-1437.

Kowalczykowski SC & Eggleston AK (1994). Annu. Rev. Biochem 63:991-1034.

Kunkel T (1985). Proc. Natl. Acad. Sci. USA 82:488-492.

Lander ES, et al. (1987). Genomics 1:174-181.

Lister C & Dean C (1993). Plant J 4:745-750.

Lord JM, Hartley MR & Roberts LM (1991). Semin Cell Biol. 2:15-22.

Mett VL, et al. (1993). Proc. Natl. Acad. Sci. USA 90:4567-4571.

Meshi T & Iwabuchi M (1995). Plant Cell Physiol. 36:1405-1420.

Mitchell AP (1995). In Microbiological Reviews, Amererican Soc for Microbiology, pp.56-70.

Napoli C, et al. (1990). Plant Cell 2:279-289.

Ohad N, et al. (1996). Proc. Natl. Acad. Sci. USA 93:5319-5324.

O'Keefe DP, et al. (1994). Plant Physiol. 105:473-482.

Onouchi H, et al. (1995). Mol. Gen. Genet. 247:653-660.

Osborne BI, et al. (1995). Plant J. 7:687-701

Puchta H, et al. (1996). Proc. Natl. Acad. Sci. USA 93:5055-5060.

Ramachandran S, et al. (1995). Curr. Opin. Genet. Dev. 4:642-646.

Riley R, et al. (1968). Genet. Res. 12:199-219.

Sambrook J, et al. (1989). Molecular Cloning. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Sato S, et al. (1995). DNA Res. 2:89-93.

Schwarzacher T & Heslop-Harrison JS (1995). In: P. E. Brandham and M. D. Bennet (editors). Kew Chromosome Conference IV, pp 407-416.

Sherman JD & Stack SUBJECT-MATTER (1992). Genome 35: 354-359.

Sealey L & Chalkley R (1987). Mol. Cell. Biol. 7:787-798.

Sears ER (1976). Annu. Rev. Genet. 10:31-51.

Sheen J, et al. (1995). Plant J. 8:777-784.

Shinohara A, et al. (1992). Cell 69:457-470.

Stone JM & Walker JC (1995). Plant Physiol. 108:451-457.

Sym M, et al. (1993). Cell 72:365-378.
Tebb G, and Mattaj IW (1989). Mol.Cell. Biol. 9:1682-1690.
Terasawa M, et al. (1995). Genes Dev. 9: 925-934.
Valvekens D, et al. (1988). Proc. Natl. Acad. Sci. USA 85:5536-5540.
Walker JE, et al. (1982) . EMBO J. 1:945-951.
Yoshida K, et al. (1995). Appl. Microbiol. Biotechnol. 44:466-472.

Claims

1. A nucleic acid molecule comprising

(a) the Arabidopis meiosis-specific DMC1 (AtDMC1) gene promoter; or

(b) a meiosis-specific promoter homologous to (a) but from another plant species; or

(c) a meiosis-specific promoter of the gene of a homologous DMC1 protein from another plant species; or

(d) a meiosis-specific promoter variant, mutant, allele or derivative of (a), (b) or (c); or

(e) a portion of (a), (b), (c) or (d) sufficient to confer meiosis-specific character to a promoter containing it.

2. A nucleic acid molecule according to claim 1 wherein the homologous DMC1 protein is encoded by a nucleic acid sequence having at least 55% homology with the nucleic acid sequence of AtDMC1.

3. A promoter comprising at least a portion of a nucleic acid molecule according to claim 1 or claim 2 sufficient to confer meiosis-specific character to the promoter.

4. A promoter according to claim 3 comprising all or part of the sequence shown in Fig. 4(A).

5. A promoter according to claim 4 wherein the sequence is derived from nucleic acid which lies 5' of nucleotide 4473 in the sequence shown in Fig. 4(A).

6. A nucleic acid construct comprising a meiosis-specific promoter according to any one of the preceding claims operably linked to a heterologous nucleic acid sequence.

7. A nucleic acid construct according to claim 6 wherein the heterologous nucleic acid sequence is a gene.

8. A nucleic acid construct according to claim 7 wherein the heterologous gene encodes a cytotoxic protein.

9. A nucleic acid vector comprising a promoter or nucleic acid construct according to any one of the preceding claims.

10. A recombinant host cell containing a promoter according to any one of claims 1 to 5, or a nucleic acid construct according to any one of claims 6 to 8 or a nucleic acid vector according to claim 9, optionally integrated into the genome of the host cell.

11. A host cell according to claim 10 being a plant cell.

12. A method of producing a plant cell of claim 11 comprising the steps of introducing a nucleic acid comprising a promoter, construct or vector into a plant cell; and causing or allowing recombination between the nucleic acid and the plant cell genome to introduce the sequence of said nucleic acid into the genome.

13. Use of a promoter, construct or vector according to any one of claims 1 to 9 in the production of a transgenic plant.

14. A plant cell comprising in its genome a promoter, construct or vector according to any one of claims 1 to 9.

15. A plant, part or propagule thereof, seed, selfed or hybrid progeny or descendant thereof comprising a plant

cell according to claim 16.

16. A method of inducing meiosis-specific expression of a nucleic acid sequence in a transgenic plant cell comprising the steps of

(a) transforming a plant cell with a nucleic acid comprising a promoter, construct or vector according to any one of claims 1 to 9 so as to introduce said nucleic acid into the genome of said plant cell so that expression of a nucleic acid sequence is regulated by meiosis-specific promoter; and optionally

(b) regenerating a plant from said plant cell.

17. A method of producing a sterile plant comprising the steps of modifying plant cells by incorporating into their genome nucleic acid comprising a modification nucleic acid sequence and a promoter, construct or vector according to any one of claims 1 to 9 so that said modification nucleic acid sequence is expressed in said meiotic cells under the control of the meiosis-specific promoter, thereby altering the meiotic cell cycle and rendering the plant sterile.

18. A method according to claim 17 wherein the construct further comprises an insertion sequence flanked by target sequences and positioned in between said nucleic acid molecule and said modification nucleic acid sequence, said target sequences being under the control of an inducible promoter such that on induction the insertion sequence is eliminated.

19. A method according to claim 17 or 18 wherein the target sequences are lox recombination sites which are activated by cre recombinase under the control of the inducible promoter.

20. A method according to any one of claims 17 to 19

wherein the modification nucleic acid sequence encodes a cytotoxic protein which leads to the ablation of meiosis cells.

21. A method of isolating DMC1 homologues comprising the steps of PCR on template nucleic acid using any one or more of the following degenerate primers

MEI1- 5' GG(N) AA(GA) GT(N) GC(N) TA(CT) AT(ACT) GA 3';

MEI4 - 5' AC(N) GC(N) AC(GA) TT(GA) AA(CT) TC(CT) TC(N) GC 3';

MEI5 - 5' GC(GA) TG(N) GC(N) A(GA)N AC(GA) TG(N) CC(N) CC 3'; and

isolating said PCR product.

22. A method according to claim 21 wherein the MEI1 and MEI5 or MEI1 and MEI4 primers are used in combination.

23. A method according to claim 21 or claim 22 wherein the MEI1 primer is 5' GGA GGG AAT GGA AAA GTG 3' and the MEI4 is 5' GCA ACG TTG AAC TCC TCT GCA AT 3'.

24. A nucleic acid molecule comprising any one of the following sequences for use as identification probes or PCR primers

MEI1- 5' GG(N) AA(GA) GT(N) GC(N) TA(CT) AT(ACT) GA 3';

MEI4- 5' AC(N) GC(N) AC(GA) TT(GA) AA(CT) TC(CT) TC(N) GC 3';

MEI5- 5' GC(GA) TG(N) GC(N) A(GA)N AC(GA) TG(N) CC(N) CC 3';

MEI1U- 5' GGA GGG AAT GGA AAA GTG 3'; and

MEI4U- 5' GCA ACG TTG AAC TCC TCT GCA AT 3'.

25. A method of transposon tagging comprising the steps of creating a construct comprising a promoter according to any one of claims 1 to 5 operably linked to a

transposase required for transposition of a transposable element; transforming a plant cell with said construct such that the transposase is driven by said promoter; and determining transposition events.

26. A nucleic acid construct for selectively expressing a nucleic acid sequence of interest comprising

a promoter according to any one of claims 1 to 5 upstream of said nucleic acid sequence of interest;

a nucleic acid insertion sequence flanked by target sequences and positioned in between said promoter and said nucleic acid sequence of interest; and

an activating gene under the control of an inducible promoter such that on induction said activating gene expresses an activator which activates said target sequences thereby eliminating said insertion sequence allowing the nucleic acid of interest to be expressed under the control of the promoter.

27. A nucleic acid construct according to claim 26 wherein the activating gene is a DNA recombinase gene.

28. A nucleic acid construct according to claim 27 wherein the activating gene is cre recombinase and the target sequences are lox recombination sites.

29. A method for selectively expressing a nucleic acid sequence of interest in a plant comprising the steps of

transforming said plant cell with a nucleic acid construct according to any one of claims 26 to 28;

regenerating a plant from said plant cell; and

inducing said inducible promoter.

30. A method of producing seeds apomictically comprising the steps of modifying plant cells by incorporating into their genome a nucleic acid construct according to any

one of claims 26 to 28 so that said modification nucleic acid sequence is expressed in said meiotic cells thereby eliminating sexually produced seeds in a plant regenerated therefrom.

31. A method according to claim 30 wherein the construct is incorporated into the genome by breeding.

32. A method according to claim 30 wherein the construct is incorporated into the genome by transformation.

33. A method of detecting apomictic mutants comprising the steps of generating a plant or plant part carrying a nucleic acid construct according to any one of claims 26 to 29; creating a variation for apomixis in accordance with a breeding programme wherein there is a natural variation for apomixis; deriving a suitable self progeny population; inducing said activating gene in the self-progeny and detecting viable seeds produced apomictically.

34. A method of detecting apomictic mutants comprising the steps of generating a plant or plant part carrying a nucleic acid construct according to any one of claims 26 to 29; mutagenising said plant or plant part in accordance with a mutagenesis programme; deriving a suitable self progeny population; inducing said activating gene in the self-progeny and detecting viable seeds produced apomictically.

35. A nucleic acid construct for selectively removing one or more transgenes from the genome of a transgeneic plant comprising

a promoter according to any one of claims 1 to 5 operably linked to an activator gene for expressing an activator; and

at least one target sequence flanking said transgene

and said target sequences being activated by said activator to eliminate said transgenes.

36. A nucleic acid construct according to claim 35 wherein the activating gene is a DNA recombinase gene.

37. A nucleic acid construct according to claim 35 or claim 36 wherein the activating gene is cre recombinase and the target sequences are lox recombination sites.

38. A method of selectively removing one or more transgenes from the genome of a transgenic plant comprising the steps of

transforming a plant cell with a construct according to any one of claims 35 to 37;

regenerating said transgenic plant from said plant cell; and

recovering progeny from meiotic reproduction of said transgenic plant

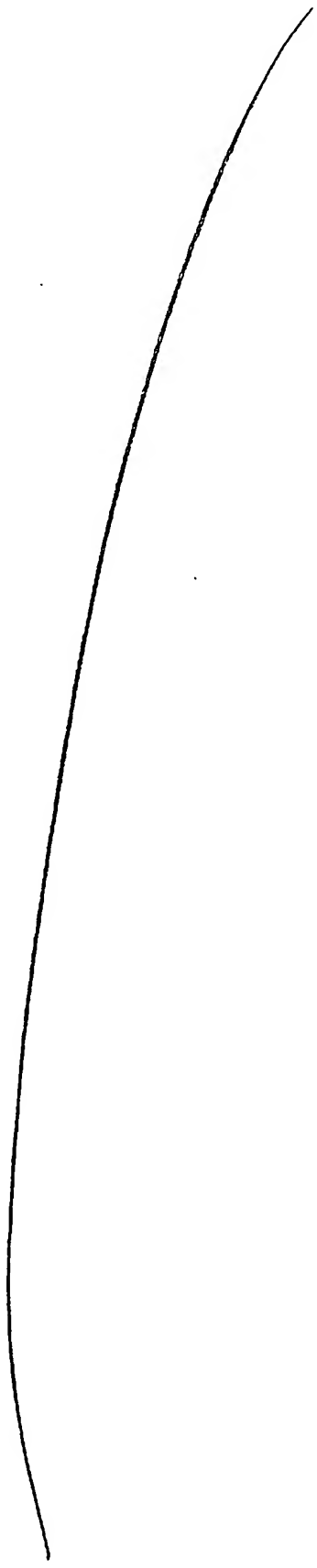
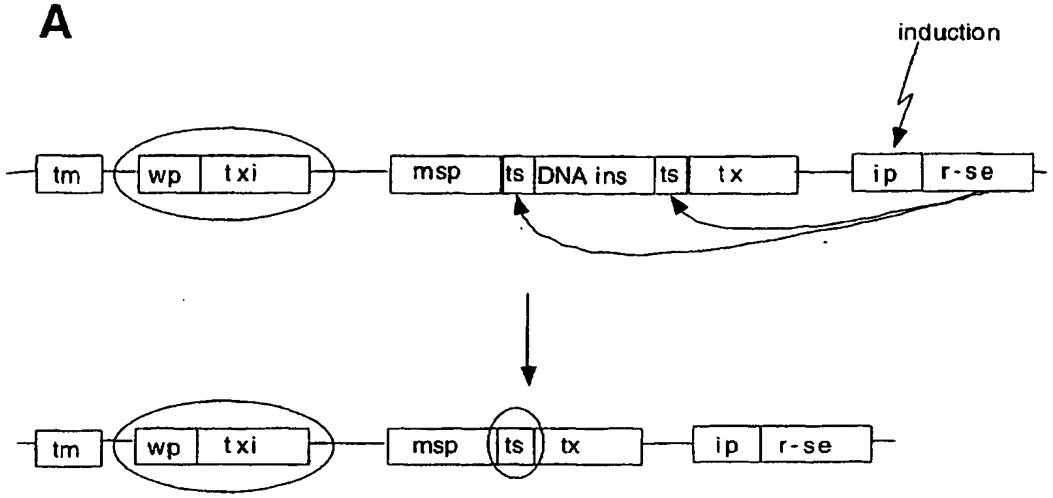
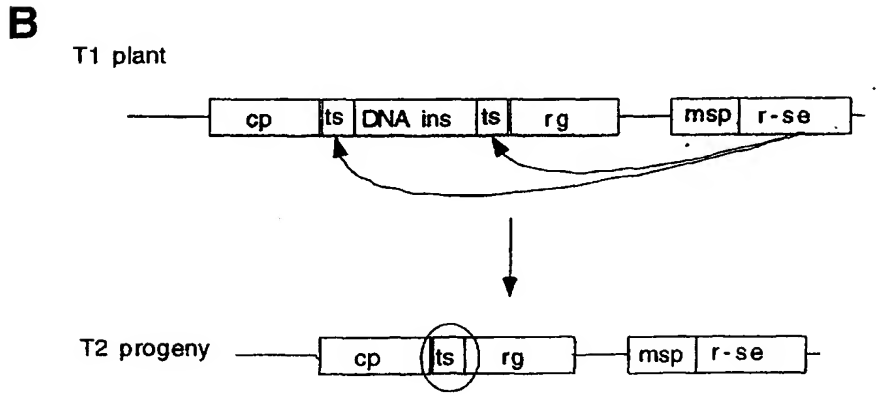


FIGURE1


A
induction
tm
wp
txi
msp
ts
DNA ins
ts
tx
ip
r-se
tm
wp
txi
msp
ts
tx
ip
r-se


B
T1 plant
cp
ts
DNA ins
ts
rg
msp
r-se
T2 progeny
cp
ts
rg
msp
r-se

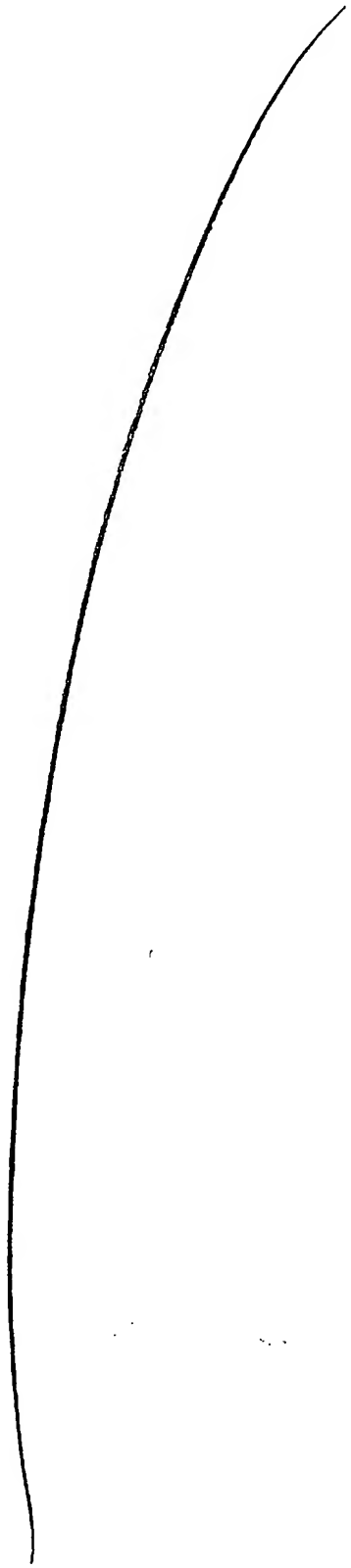
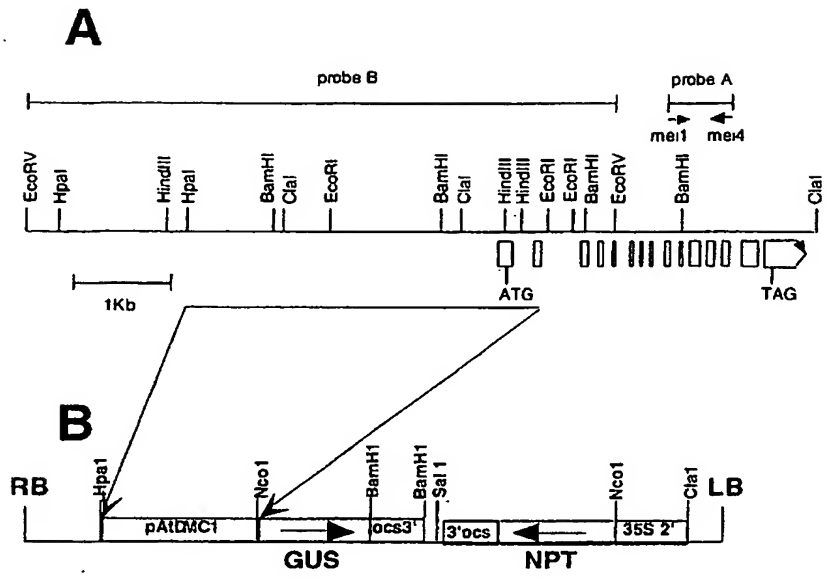
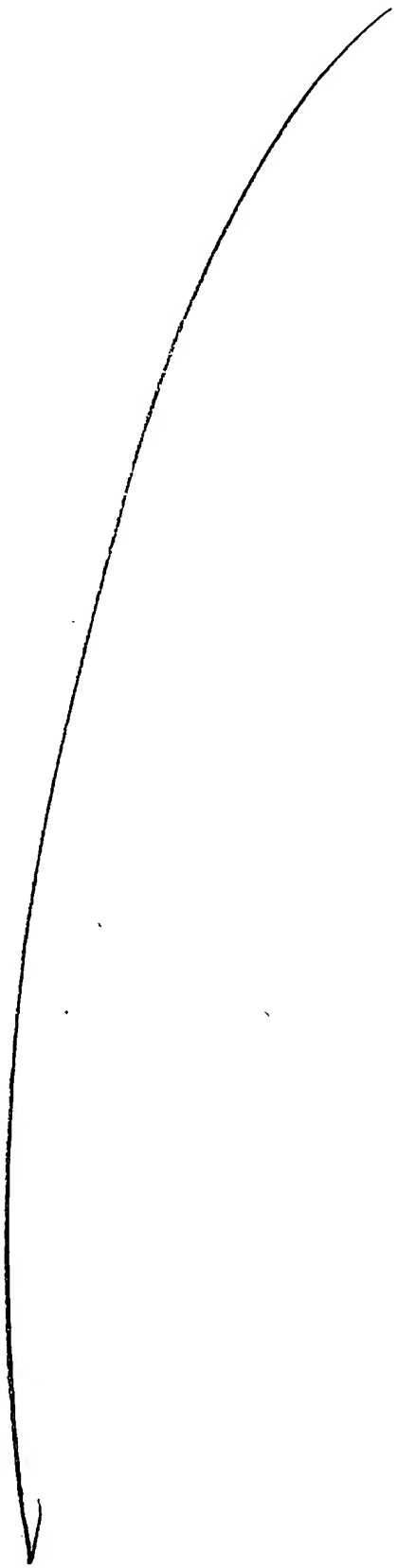
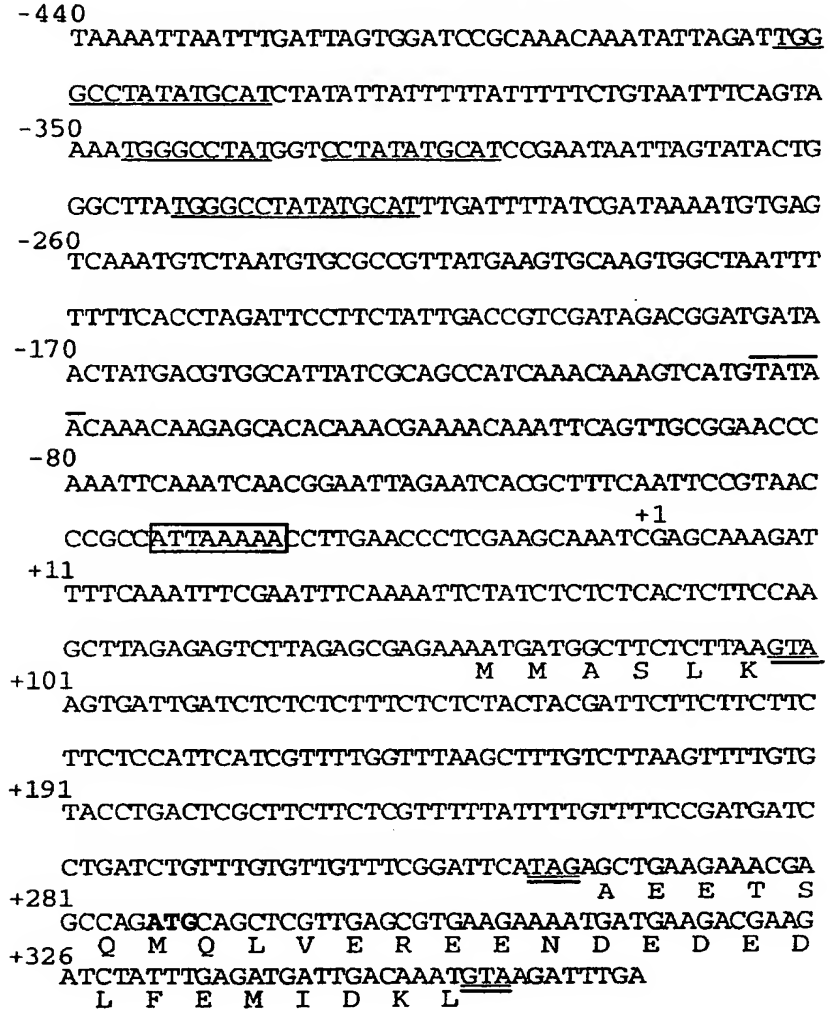


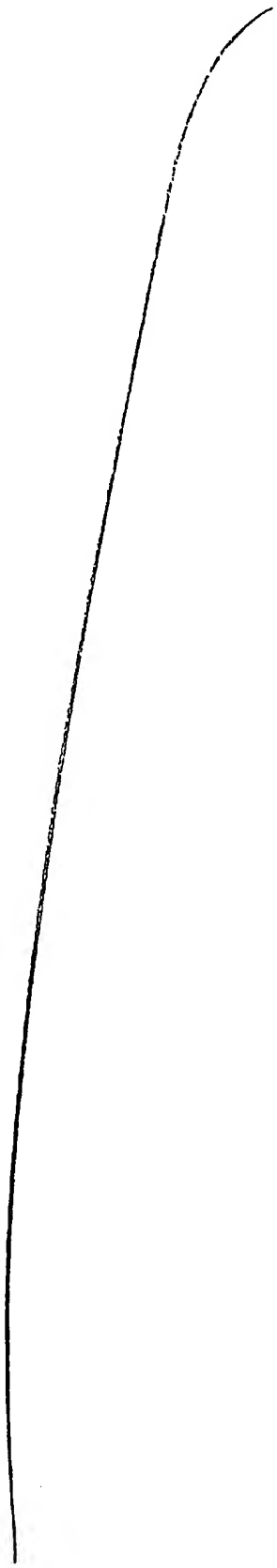
FIGURE 2


A
probe B
probe A
mei1
mei4
EcoRV
HpaI
HindIII
HpaI
BamHI
ClaI
EcoRI
BamHI
ClaI
HindIII
HindIII
EcoRI
EcoRI
BamHI
EcoRV
BamHI
ClaI
1Kb
ATG
TAG
B
RB
Hpa1
Nco1
BamH1
BamH1
Sal 1
Nco1
Cla1
LB
pAtDMC1
ocs3'
3'ocs
35S 2'
GUS
NPT

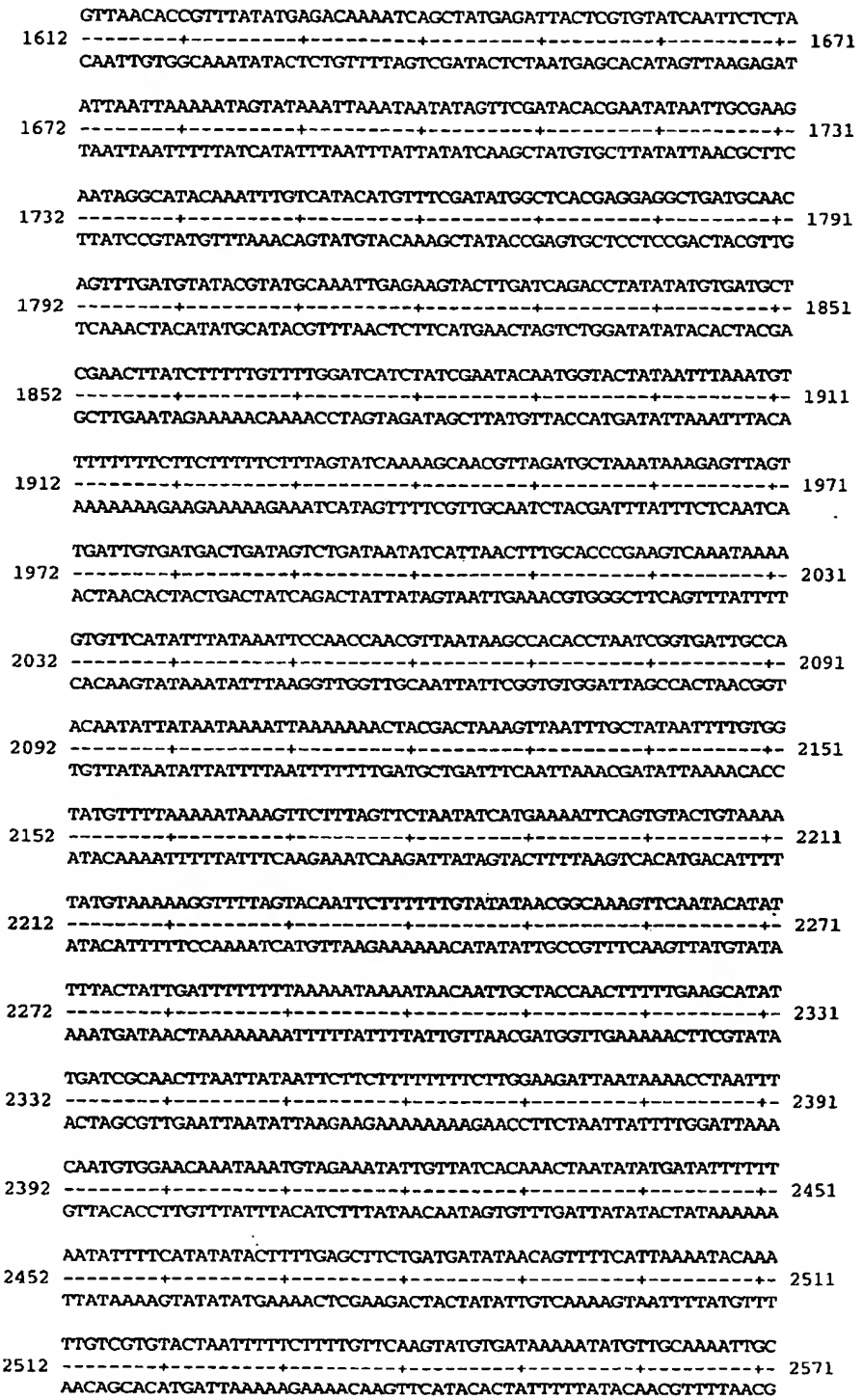


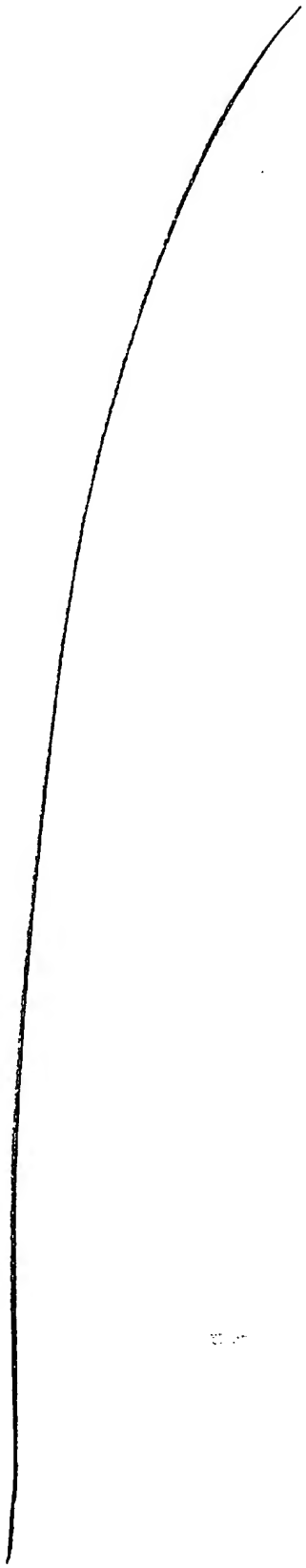
# FIGURE3.


-440
TAAAATTAATTTGATTAGTGGATCCGCAAACAAATATTAGATTGG
GCCTATATGCATCTATATTATTTTTATTTTTCTGTAATTTCAGTA
-350
AAATGGGCCTATGGTCCTATATGCATCCGAATAATTAGTATACTG
GGCTTATGGGCCTATATGCATTTGATTTTATCGATAAAATGTGAG
-260
TCAAATGTCTAATGTGCGCCGTTATGAAGTGCAAGTGGCTAATTT
TTTTCACCTAGATTCCTTCTATTGACCGTCGATAGACGGATGATA
-170
ACTATGACGTGGCATTATCGCAGCCATCAAACAAAGTCATGTATA
ACAAACAAGAGCACACAAACGAAAACAAATTCAGTTGCGGAACCC
-80
AAATTCAAATCAACGGAATTAGAATCACGCTTTCAATTCCGTAAC
+1
CCGCCATTAAAAACCTTGAACCCTCGAAGCAAATCGAGCAAAGAT
+11
TTTCAAATTTCGAATTTCAAAATTCTATCTCTCTCACTCTTCCAA
GCTTAGAGAGTCTTAGAGCGAGAAAATGATGGCTTCTCTTAAGTA
M M A S L K
+101
AGTGATTGATCTCTCTCTTTCTCTCTACTACGATTCTTCTTCTTC
TTCTCCATTCATCGTTTTGGTTTAAGCTTTGTCTTAAGTTTTGTG
+191
TACCTGACTCGCTTCTTCTCGTTTTTATTTTGTTTTCCGATGATC
CTGATCTGTTTGTGTTGTTTCGGATTCATAGAGCTGAAGAAACGA
A E E T S
+281
GCCAGATGCAGCTCGTTGAGCGTGAAGAAAATGATGAAGACGAAG
Q M Q L V E R E E N D E D E D
+326
ATCTATTTGAGATGATTGACAAATGTAAGATTTGA
L F E M I D K L

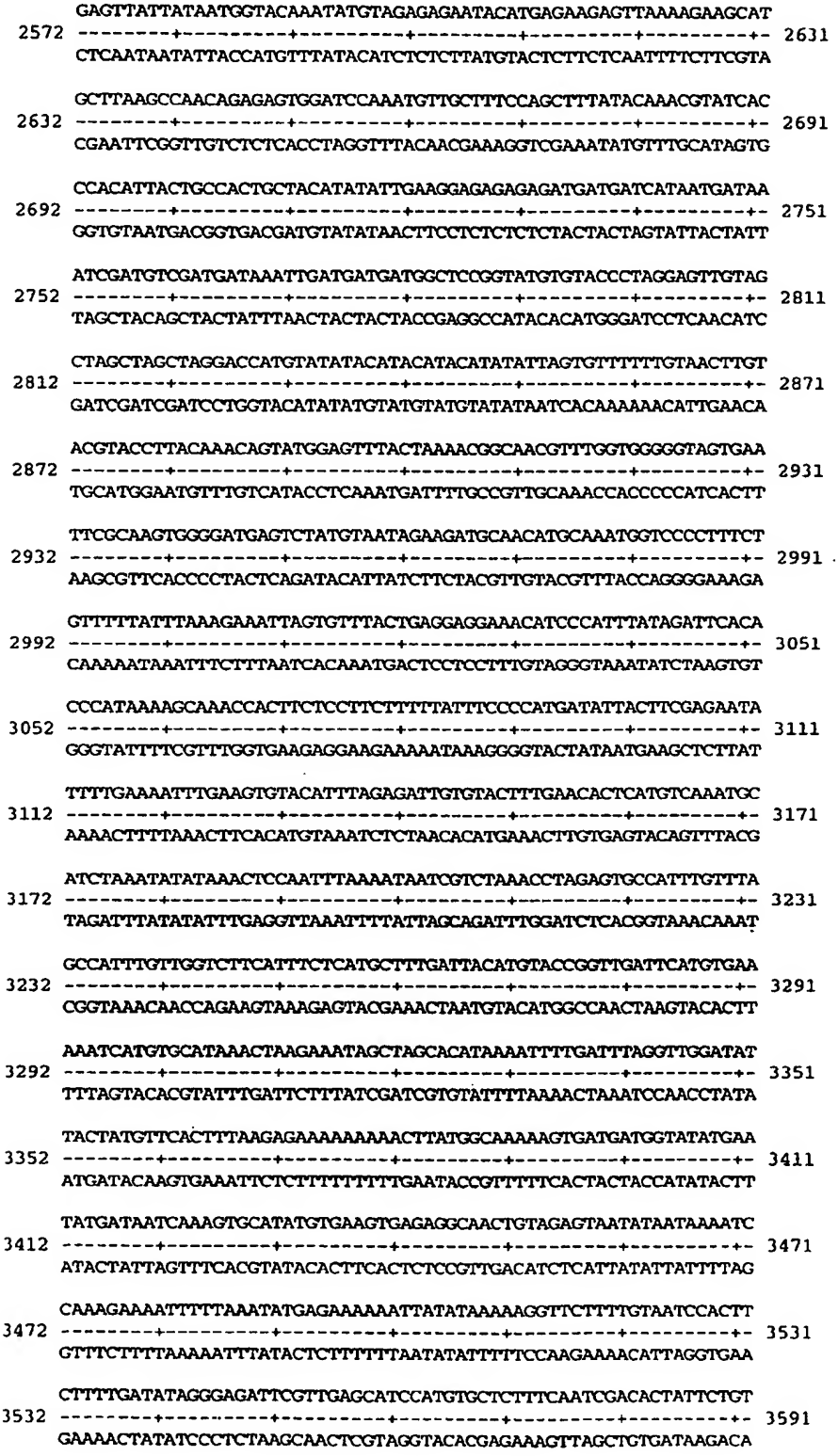


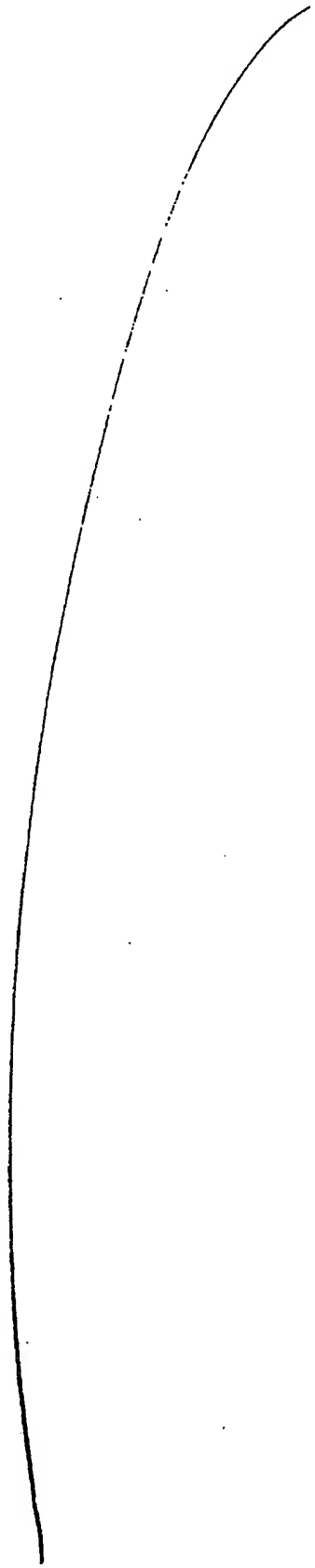
# FIGURE4 (A)


GTTAACACCGTTTATATGAGACAAAATCAGCTATGAGATTACTCGTGTATCAATTCTCTA
1612 ---------+---------+---------+---------+---------+---------+- 1671
CAATTGTGGCAAATATACTCTGTTTTAGTCGATACTCTAATGAGCACATAGTTAAGAGAT
ATTAATTAAAAATAGTATAAATTAAATAATATAGTTCGATACACGAATATAATTGCGAAG
1672 ---------+---------+---------+---------+---------+---------+- 1731
TAATTAATTTTTATCATATTTAATTTATTATATCAAGCTATGTGCTTATATTAACGCTTC
AATAGGCATACAAATTTGTCATACATGTTTCGATATGGCTCACGAGGAGGCTGATGCAAC
1732 ---------+---------+---------+---------+---------+---------+- 1791
TTATCCGTATGTTTAAACAGTATGTACAAAGCTATACCGAGTGCTCCTCCGACTACGTTG
AGTTTGATGTATACGTATGCAAATTGAGAAGTACTTGATCAGACCTATATATGTGATGCT
1792 ---------+---------+---------+---------+---------+---------+- 1851
TCAAACTACATATGCATACGTTTAACTCTTCATGAACTAGTCTGGATATATACACTACGA
CGAACTTATCTTTTTGTTTTGGATCATCTATCGAATACAATGGTACTATAATTTAAATGT
1852 ---------+---------+---------+---------+---------+---------+- 1911
GCTTGAATAGAAAAACAAAACCTAGTAGATAGCTTATGTTACCATGATATTAAATTTACA
TTTTTTTCTTCTTTTTCTTTAGTATCAAAAGCAACGTTAGATGCTAAATAAAGAGTTAGT
1912 ---------+---------+---------+---------+---------+---------+- 1971
AAAAAAAGAAGAAAAAGAAATCATAGTTTTCGTTGCAATCTACGATTTATTTCTCAATCA
TGATTGTGATGACTGATAGTCTGATAATATCATTAACTTTGCACCCGAAGTCAAATAAAA
1972 ---------+---------+---------+---------+---------+---------+- 2031
ACTAACACTACTGACTATCAGACTATTATAGTAATTGAAACGTGGGCTTCAGTTTATTTT
GTGTTCATATTTATAAATTCCAACCAACGTTAATAAGCCACACCTAATCGGTGATTGCCA
2032 ---------+---------+---------+---------+---------+---------+- 2091
CACAAGTATAAATATTTAAGGTTGGTTGCAATTATTCGGTGTGGATTAGCCACTAACGGT
ACAATATTATAATAAAATTAAAAAAACTACGACTAAAGTTAATTTGCTATAATTTTGTGG
2092 ---------+---------+---------+---------+---------+---------+- 2151
TGTTATAATATTATTTTAATTTTTTTGATGCTGATTTCAATTAAACGATATTAAAACACC
TATGTTTTAAAAATAAAGTTCTTTAGTTCTAATATCATGAAAATTCAGTGTACTGTAAAA
2152 ---------+---------+---------+---------+---------+---------+- 2211
ATACAAAATTTTTATTTCAAGAAATCAAGATTATAGTACTTTTAAGTCACATGACATTTT
TATGTAAAAAGGTTTTAGTACAATTCTTTTTTGTATATAACGGCAAAGTTCAATACATAT
2212 ---------+---------+---------+---------+---------+---------+- 2271
ATACATTTTTCCAAAATCATGTTAAGAAAAAACATATATTGCCGTTTCAAGTTATGTATA
TTTACTATTGATTTTTTTTAAAAATAAAATAACAATTGCTACCAACTTTTTGAAGCATAT
2272 ---------+---------+---------+---------+---------+---------+- 2331
AAATGATAACTAAAAAAAAATTTTTATTTTATTGTTAACGATGGTTGAAAAACTTCGTATA
TGATCGCAACTTAATTATAATTCTTCTTTTTTTTCTTGGAAGATTAATAAAACCTAATTT
2332 ---------+---------+---------+---------+---------+---------+- 2391
ACTAGCGTTGAATTAATATTAAGAAGAAAAAAAAGAACCTTCTAATTATTTTGGATTAAA
CAATGTGGAACAAATAAATGTAGAAATATTGTTATCACAAACTAATATATGATATTTTTT
2392 ---------+---------+---------+---------+---------+---------+- 2451
GTTACACCTTGTTTATTTACATCTTTATAACAATAGTGTTTGATTATATACTATAAAAAA
AATATTTTCATATATACTTTTGAGCTTCTGATGATATAACAGTTTTCATTAAAATACAAA
2452 ---------+---------+---------+---------+---------+---------+- 2511
TTATAAAAGTATATATGAAAACTCGAAGACTACTATATTGTCAAAAGTAATTTTATGTTT
TTGTCGTGTACTAATTTTTCTTTTGTTCAAGTATGTGATAAAAATATGTTGCAAAATTGC
2512 ---------+---------+---------+---------+---------+---------+- 2571
AACAGCACATGATTAAAAAGAAAACAAGTTCATACACTATTTTTATACAACGTTTTAACG

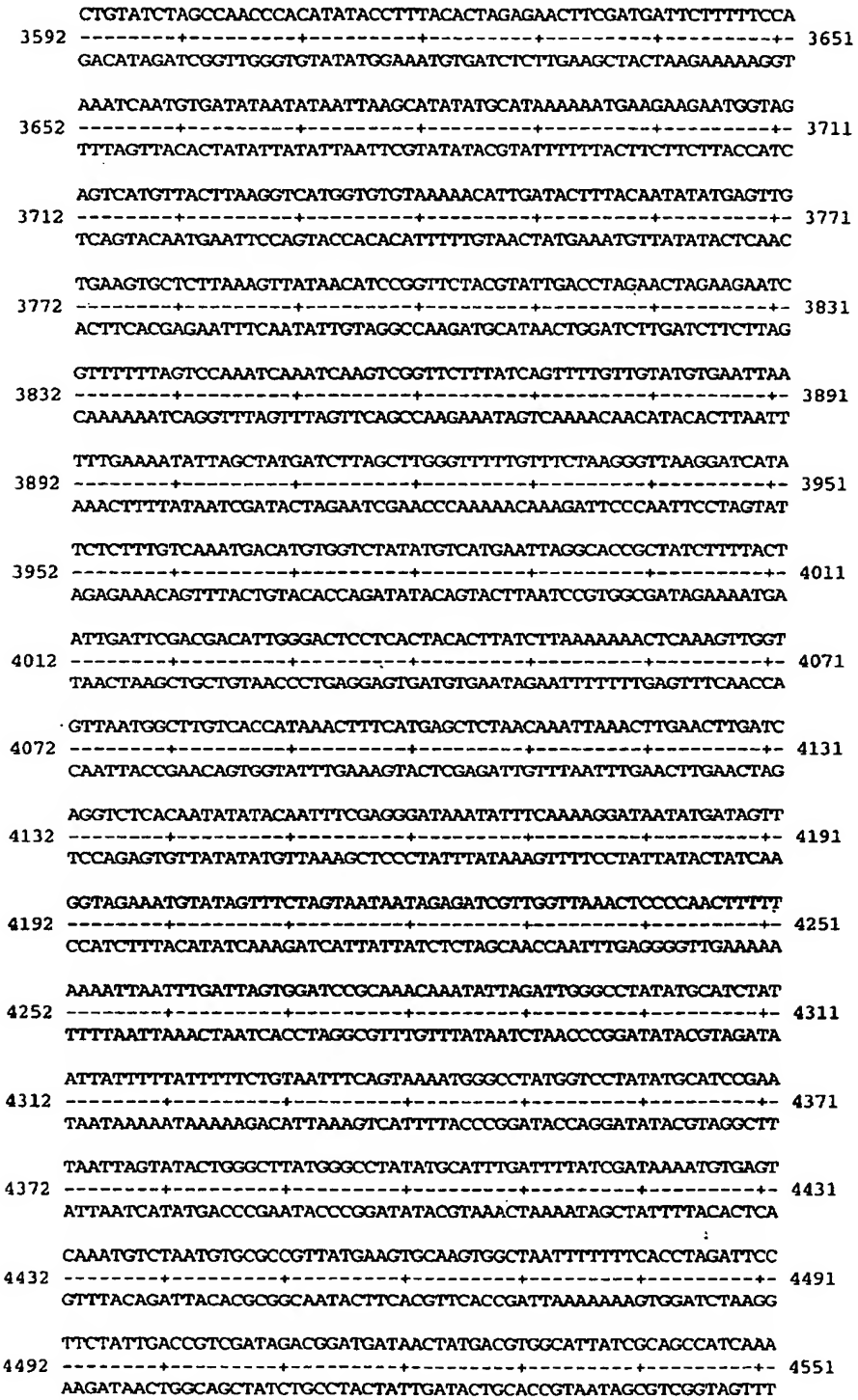


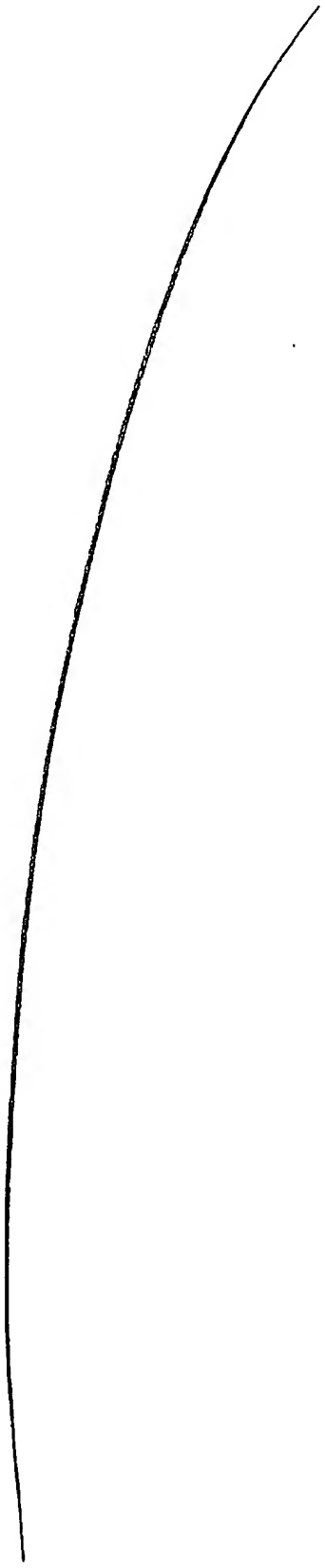
# FIGURE4(A)continued





# FIGURE4(A)continued



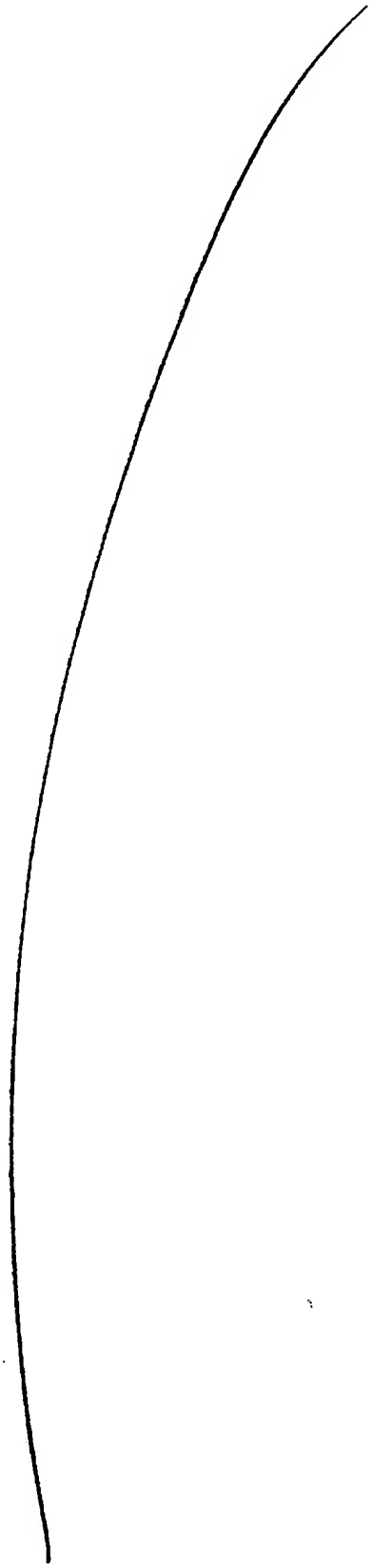


# FIGURE4(A)continued

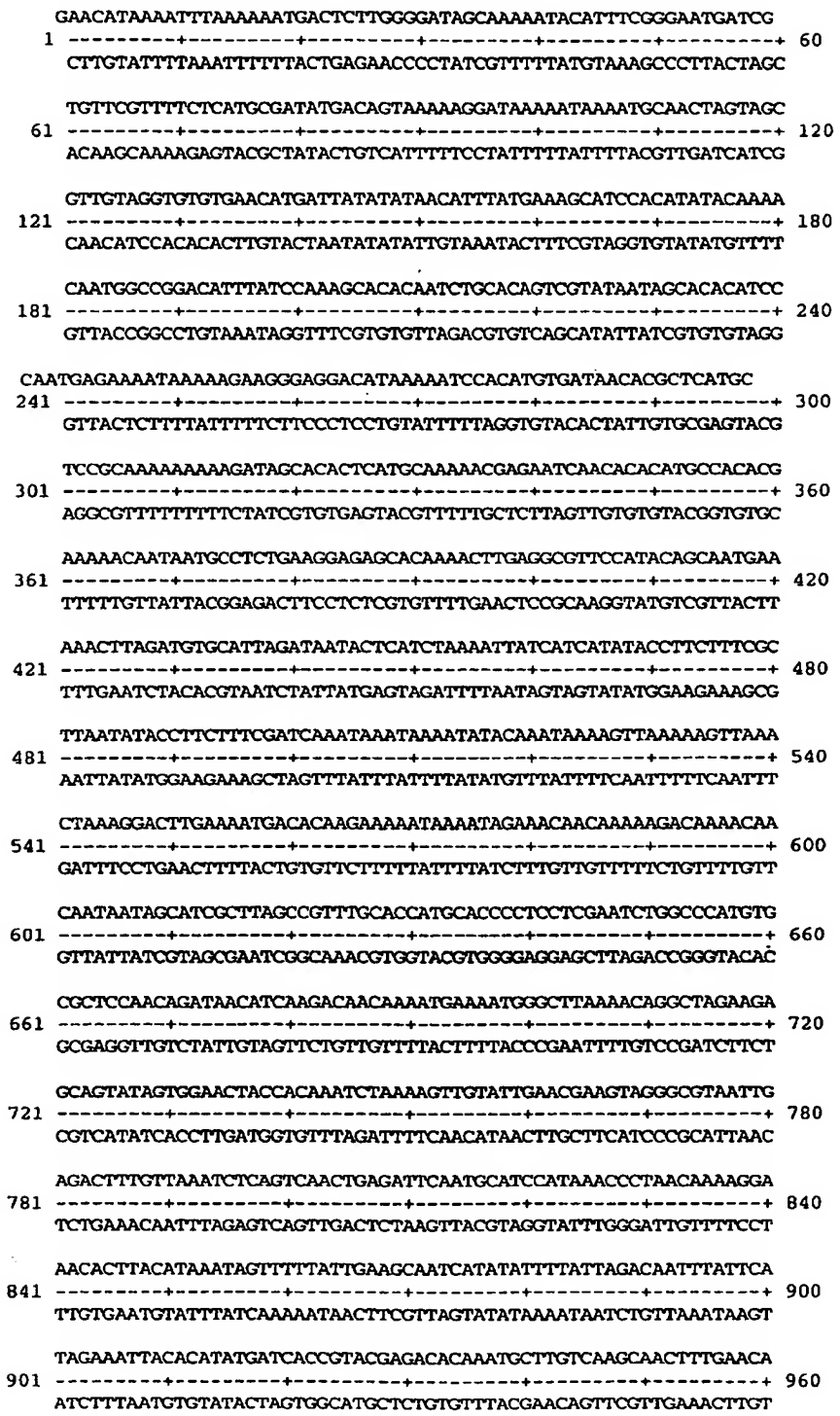
```
     CAAAGTCATGTATAACAAACAAGAGCACACAAACGAAAACAAATTCAGTTGCGGAACCCA
4552 --------+---------+---------+---------+---------+---------+-- 4611
     GTTTCAGTACATATTGTTTGTTCTCGTGTGTTTGCTTTTGTTTAAGTCAACGCCTTGGGT

     AATTCAAATCAACGGAATTAGAATCACGCTTTCAATTCCGTAACCCGCCATTAAAAACCT
4612 --------+---------+---------+---------+---------+---------+-- 4671
     TTAAGTTTAGTTGCCTTAATCTTAGTGCGAAAGTTAAGGCATTGGGCGGTAATTTTTGGA

     TGAACCCTCGAAGCAAATCGAGCAAAGATTTTCAAATTTCGAATTTCAAAATTCTATCTC
4672 --------+---------+---------+---------+---------+---------+-- 4731
     ACTTGGGAGCTTCGTTTAGCTCGTTTCTAAAAGTTTAAAGCTTAAAGTTTTAAGATAGAG

     TCTCACTCTTCCAAGCTTAGAGAGTCTTAGAGCGAGAAAATGATGGCT
4732 --------+---------+---------+---------+--------- 4779
     AGAGTGAGAAGGTTCGAATCTCTCAGAATCTCGCTCTTTTACTACCGA
```

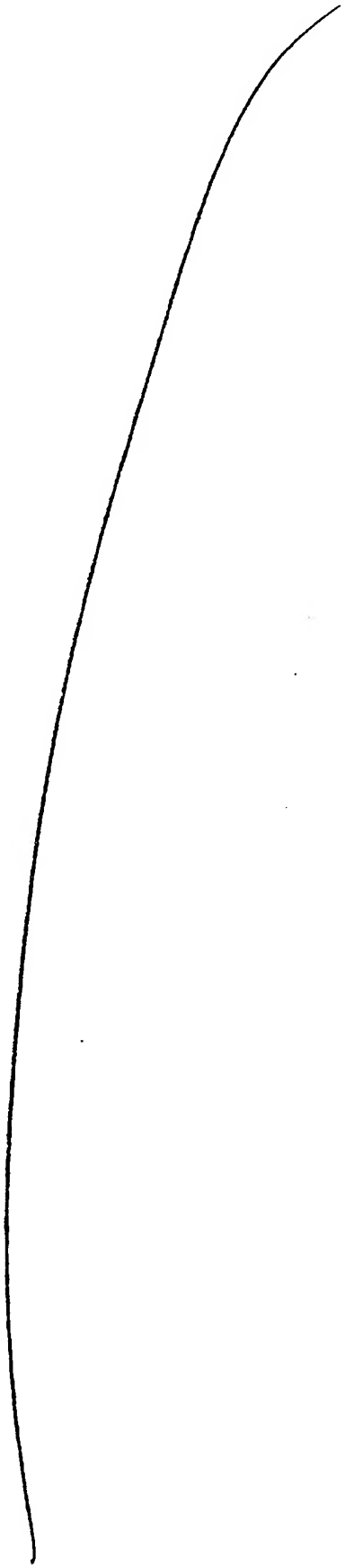
# FIGURE4(B)

```
4473 TTTTTTTCACCTAGATTCCTTCTATTGACCGTCGATAGACGGATGATAAC 4522
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1120 ttttttcacctagattccttctattgaccgtcgatagacggatgataac 1169

4523 TATGACGTGGCATTATCGCAGCCATCAAACAAAGTCATGTATAACAAACA 4572
     ||||||||||||||||||||||||||||||||||||||||||||||||| |
1170 tatgacgtggcattatcgcagccatcaaacaaagtcatgtataacaaaga 1219

4573 AGAGCACACAAACGAAAACAAATTCAGTTGCGGAACCCAAATTCAAATCA 4622
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1220 agagcacacaaacgaaaacaaattcagttgcggaacccaaattcaaatca 1269

4623 AC.GGAATTAGAATCACGCTTTCAATTCCGTAACCCGCCATTAAAAACCT 4671
     || ||||||||||||||||||||||||||||||||||||||||||||||||
1270 acgggaattagaatcacgctttcaattccgtaacccgccattaaaaacct 1319

4672 TGAACCCTCGAAGCAAATC 4690
     ||||||||||||||||||
1320 tgaaccctcgaagcaaatc 1338
```



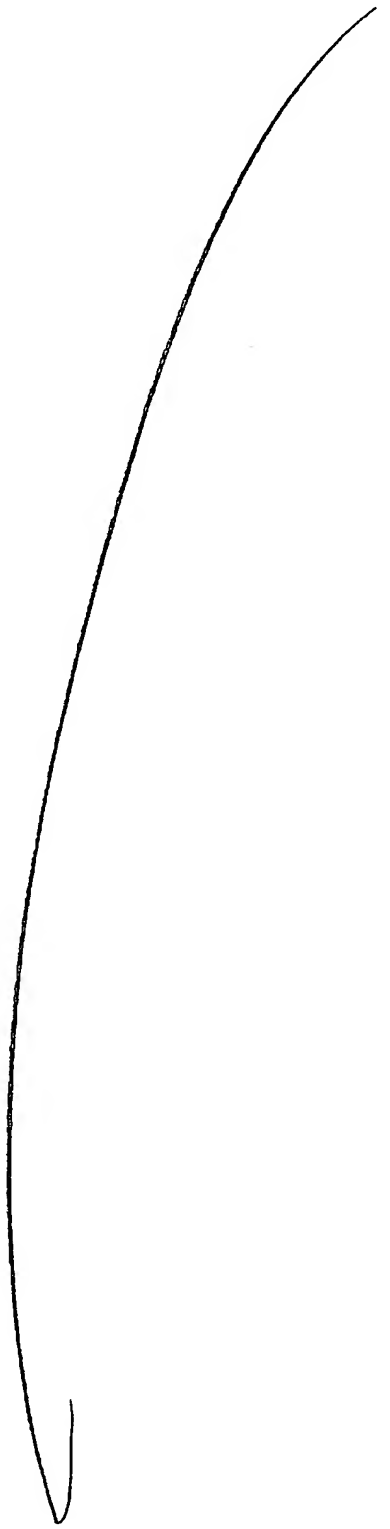
# FIGURE 5(A)


GAACATAAAATTTAAAAAATGACTCTTGGGGATAGCAAAAATACATTTCGGGAATGATCG
1 ---------+---------+---------+---------+---------+---------+ 60
CTTGTATTTTAAATTTTTTACTGAGAACCCCTATCGTTTTTATGTAAAGCCCTTACTAGC
TGTTCGTTTTCTCATGCGATATGACAGTAAAAAGGATAAAAATAAAATGCAACTAGTAGC
61 ---------+---------+---------+---------+---------+---------+ 120
ACAAGCAAAAGAGTACGCTATACTGTCATTTTTCCTATTTTTATTTTACGTTGATCATCG
GTTGTAGGTGTGTGAACATGATTATATATAACATTTATGAAAGCATCCACATATACAAAA
121 ---------+---------+---------+---------+---------+---------+ 180
CAACATCCACACACTTGTACTAATATATATTGTAAATACTTTCGTAGGTGTATATGTTTT
CAATGGCCGGACATTTATCCAAAGCACACAATCTGCACAGTCGTATAATAGCACACATCC
181 ---------+---------+---------+---------+---------+---------+ 240
GTTACCGGCCTGTAAATAGGTTTCGTGTGTTAGACGTGTCAGCATATTATCGTGTGTAGG
CAATGAGAAAATAAAAAGAAGGGAGGACATAAAAATCCACATGTGATAACACGCTCATGC
241 ---------+---------+---------+---------+---------+---------+ 300
GTTACTCTTTTATTTTTCTTCCCTCCTGTATTTTTAGGTGTACACTATTGTGCGAGTACG
TCCGCAAAAAAAAAAGATAGCACACTCATGCAAAAACGAGAATCAACACACATGCCACACG
301 ---------+---------+---------+---------+---------+---------+ 360
AGGCGTTTTTTTTTCTATCGTGTGAGTACGTTTTTGCTCTTAGTTGTGTGTACGGTGTGC
AAAAACAATAATGCCTCTGAAGGAGAGCACAAAACTTGAGGCGTTCCATACAGCAATGAA
361 ---------+---------+---------+---------+---------+---------+ 420
TTTTTGTTATTACGGAGACTTCCTCTCGTGTTTTGAACTCCGCAAGGTATGTCGTTACTT
AAACTTAGATGTGCATTAGATAATACTCATCTAAAATTATCATCATATACCTTCTTTCGC
421 ---------+---------+---------+---------+---------+---------+ 480
TTTGAATCTACACGTAATCTATTATGAGTAGATTTTAATAGTAGTATATGGAAGAAAGCG
TTAATATACCTTCTTTCGATCAAATAAATAAAATATACAAATAAAAGTTAAAAAGTTAAA
481 ---------+---------+---------+---------+---------+---------+ 540
AATTATATGGAAGAAAGCTAGTTTATTTATTTTATATGTTTATTTTCAATTTTTCAATTT
CTAAAGGACTTGAAAATGACACAAGAAAAATAAAATAGAAACAACAAAAAGACAAAACAA
541 ---------+---------+---------+---------+---------+---------+ 600
GATTTCCTGAACTTTTACTGTGTTCTTTTTATTTTATCTTTGTTGTTTTTCTGTTTTGTT
CAATAATAGCATCGCTTAGCCGTTTGCACCATGCACCCCTCCTCGAATCTGGCCCATGTG
601 ---------+---------+---------+---------+---------+---------+ 660
GTTATTATCGTAGCGAATCGGCAAACGTGGTACGTGGGGAGGAGCTTAGACCGGGTACAC
CGCTCCAACAGATAACATCAAGACAACAAAATGAAAATGGGCTTAAAACAGGCTAGAAGA
661 ---------+---------+---------+---------+---------+---------+ 720
GCGAGGTTGTCTATTGTAGTTCTGTTGTTTTACTTTTACCCGAATTTTGTCCGATCTTCT
GCAGTATAGTGGAACTACCACAAATCTAAAAGTTGTATTGAACGAAGTAGGGCGTAATTG
721 ---------+---------+---------+---------+---------+---------+ 780
CGTCATATCACCTTGATGGTGTTTAGATTTTCAACATAACTTGCTTCATCCCGCATTAAC
AGACTTTGTTAAATCTCAGTCAACTGAGATTCAATGCATCCATAAACCCTAACAAAAGGA
781 ---------+---------+---------+---------+---------+---------+ 840
TCTGAAACAATTTAGAGTCAGTTGACTCTAAGTTACGTAGGTATTTGGGATTGTTTTCCT
AACACTTACATAAATAGTTTTTATTGAAGCAATCATATATTTTATTAGACAATTTATTCA
841 ---------+---------+---------+---------+---------+---------+ 900
TTGTGAATGTATTTATCAAAAATAACTTCGTTAGTATATAAAATAATCTGTTAAATAAGT
TAGAAATTACACATATGATCACCGTACGAGACACAAATGCTTGTCAAGCAACTTTGAACA
901 ---------+---------+---------+---------+---------+---------+ 960
ATCTTTAATGTGTATACTAGTGGCATGCTCTGTGTTTACGAACAGTTCGTTGAAACTTGT

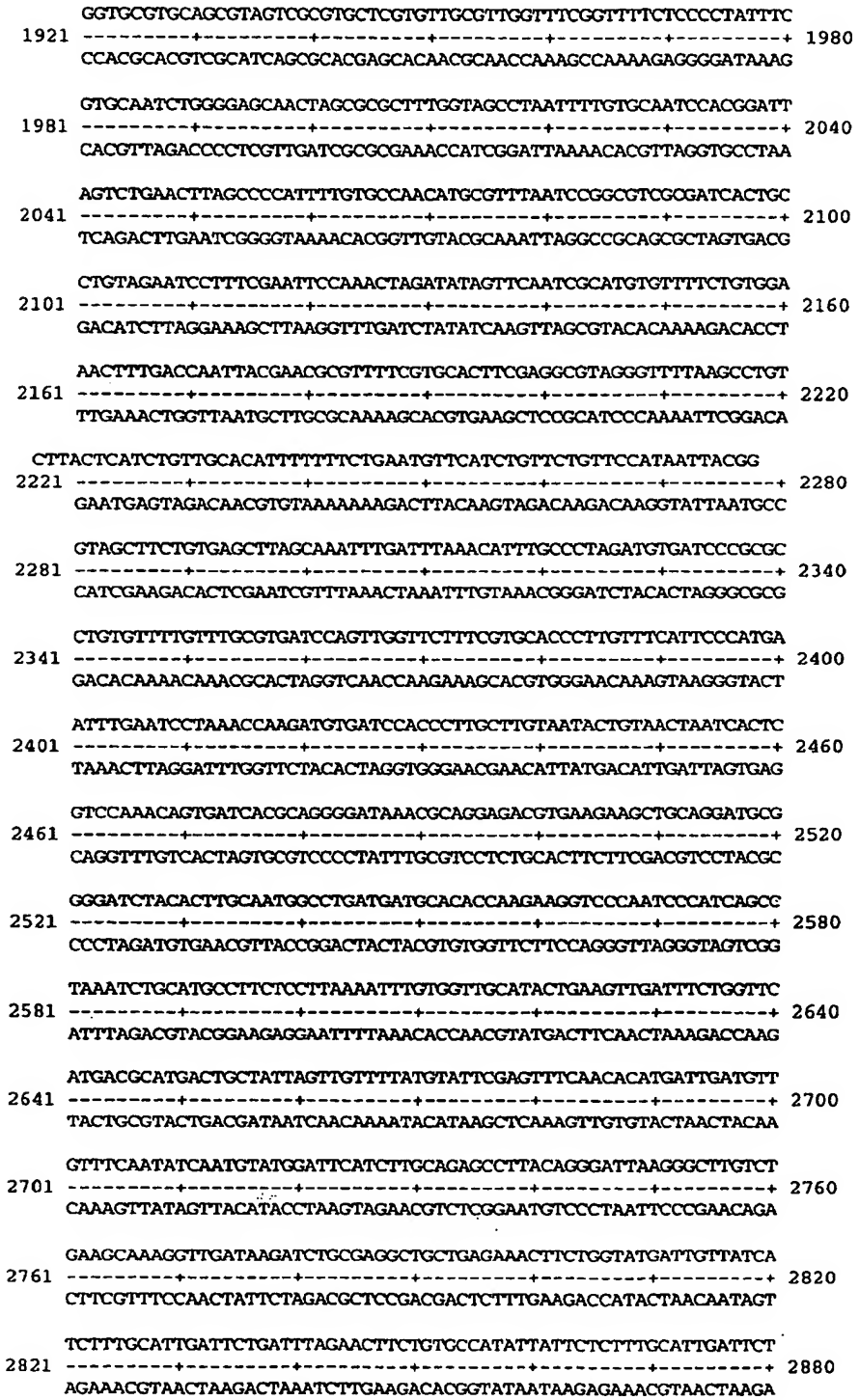


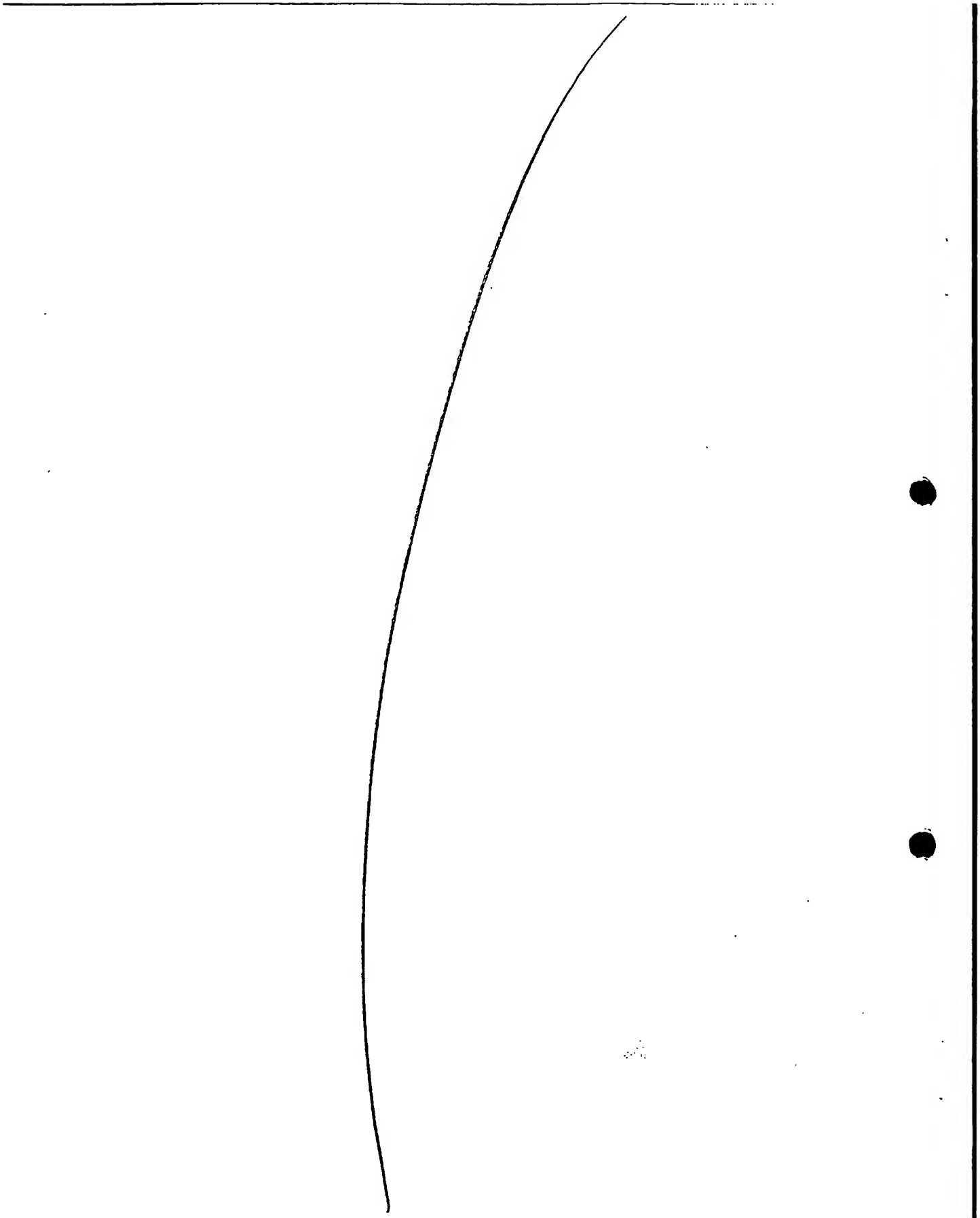
# FIGURE 5(A) continued

```
                AAAAAATTCACAAGAAGTAAAACTACTATTCCATCCTCACAGAAATTCATGCCTCGATGG
      961 ---------+---------+---------+---------+---------+---------+ 1020
                TTTTTTAAGTGTTCTTCATTTTGATGATAAGGTAGGAGTGTCTTTAAGTACGGAGCTACC

                TGAGCAGTCCACAAGCGTTTTTCACCGTCACCATGACAGCACCGGAAAAAAAAAAGCAATG
      1021 ---------+---------+---------+---------+---------+---------+ 1080
                ACTCGTCAGGTGTTCGCAAAAAGTGGCAGTGGTACTGTCGTGGCCTTTTTTTTTCGTTAC

                TCATAGCAAGATCTCGGAAAGCACCACCACACAGTAATGATATGTTAAACCCATAAAACA
      1081 ---------+---------+---------+---------+---------+---------+ 1140
                AGTATCGTTCTAGAGCCTTTCGTGGTGGTGTGTCATTACTATACAATTTGGGTATTTTGT

                GGTCATGGTAAGCTACGAGACCAGGTCTTTGTGGCATGTAGGCCGGGGATCCTCCTCGTG
      1141 ---------+---------+---------+---------+---------+---------+ 1200
                CCAGTACCATTCGATGCTCTGGTCCAGAAACACCGTACATCCGGCCCCTAGGAGGAGCAC

                GTCACCAGATCTCGATGGAGTTAACTTCATTCGATGTAGTCGAAGAAGAGGAACGTCAGC
      1201 ---------+---------+---------+---------+---------+---------+ 1260
                CAGTGGTCTAGAGCTACCTCAATTGAAGTAAGCTACATCAGCTTCTTCTCCTTGCAGTCG

                GCCCCGTTCAAAATTTTGTATTTCGAATGGGCCTAGCTGGCAGCAGCAAAGAACGTTGGG
      1261 ---------+---------+---------+---------+---------+---------+ 1320
                CGGGGCAAGTTTTAAAACATAAAGCTTACCCGGATCGACCGTCGTCGTTTCTTGCAACCC

                CTTGGTGTTACGATCCACAAGCACAAGTTCCTCGCGGGGGCCCCACAAATACCCCCACCT
      1321 ---------+---------+---------+---------+---------+---------+ 1380
                GAACCACAATGCTAGGTGTTCGTGTTCAAGGAGCGCCCCCGGGGTGTTTATGGGGGTGGA

                CGCAGCCTCAGTGATCCACCCGCACCTGCGCAACAAACCTCCCCTCTCCCCCAGCCTCCG
      1381 ---------+---------+---------+---------+---------+---------+ 1440
                GCGTCGGAGTCACTAGGTGGGCGTGGACGCGTTGTTTGGAGGGGAGAGGGGGTCGGAGGC

                CCCCAACTTCCTTCTCCAAGCCGCACGCGAGCCTCCGCCTCGTTACAGGTCTCGTCCTCC
      1441 ---------+---------+---------+---------+---------+---------+ 1500
                GGGGTTGAAGGAAGAGGTTCGGCGTGCGCTCGGAGGCGGAGCAATGTCCAGAGCAGGAGG

                TCCACCCGCTTCCACTCCACCCGTCTTCGCTCGTTGTGTCCCCGTGCTCGCTCCTAACAC
      1501 ---------+---------+---------+---------+---------+---------+ 1560
                AGGTGGGCGAAGGTGAGGTGGGCAGAAGCGAGCAACACAGGGGCACGAGCGAGGATTGTG

                CTTCTCCTATGAGGCTGTGACGCAGGCGCTCACGCGGGATGGCGCCGTCCAAGCAGTACG
      1561 ---------+---------+---------+---------+---------+---------+ 1620
                GAAGAGGATACTCCGACACTGCGTCCGCGAGTGCGCCCTACCGCGGCAGGTTCGTCATGC

                ACGAGGGCGGGCAGCTCCAGCTCATGGAGGCCGACCGGGTCGAGGAGGAGGAGGAGTGCT
      1621 ---------+---------+---------+---------+---------+---------+ 1680
                TGCTCCCGCCCGTCGAGGTCGAGTACCTCCGGCTGGCCCAGCTCCTCCTCCTCCTCACGA

                TCGAGTCCATTGACAAGTGTACGTTCGACGCCTCCTACCCCTCTCCTCTGCAAACCCTCC
      1681 ---------+---------+---------+---------+---------+---------+ 1740
                AGCTCAGGTAACTGTTCACATGCAAGCTGCGGAGGATGGGGAGAGGAGACGTTTGGGAGG

                CGCCGCCCCCGTCTCCGGCGTGCGTCGCGCGCGTGTCGGCGTGTGGGGGTTAGGCCTGCT
      1741 ---------+---------+---------+---------+---------+---------+ 1800
                GCGGCGGGGGCAGAGGCCGCACGCAGCGCGCGCACAGCCGCACACCCCCAATCCGGACGA

                CGCCGTTGCGTTCCGGGTGCCTCCGCCTCTCTAAGTTCGCGCGTTTCGGTTGCAATTTCG
      1801 ---------+---------+---------+---------+---------+---------+ 1860
                GCGGCAACGCAAGGCCCACGGAGGCGGAGAGATTCAAGCGCGCAAAGCCAACGTTAAAGC

                TGCTGTTTGGAGGTAGACTTGGTGCGGATTTTGCTTAGCCTCCACATTTGGTTGGTTTTT
      1861 ---------+---------+---------+---------+---------+---------+ 1920

                ACGACAAACCTCCATCTGAACCACGCCTAAAACGAATCGGAGGTGTAAACCAACCAAAAA
```

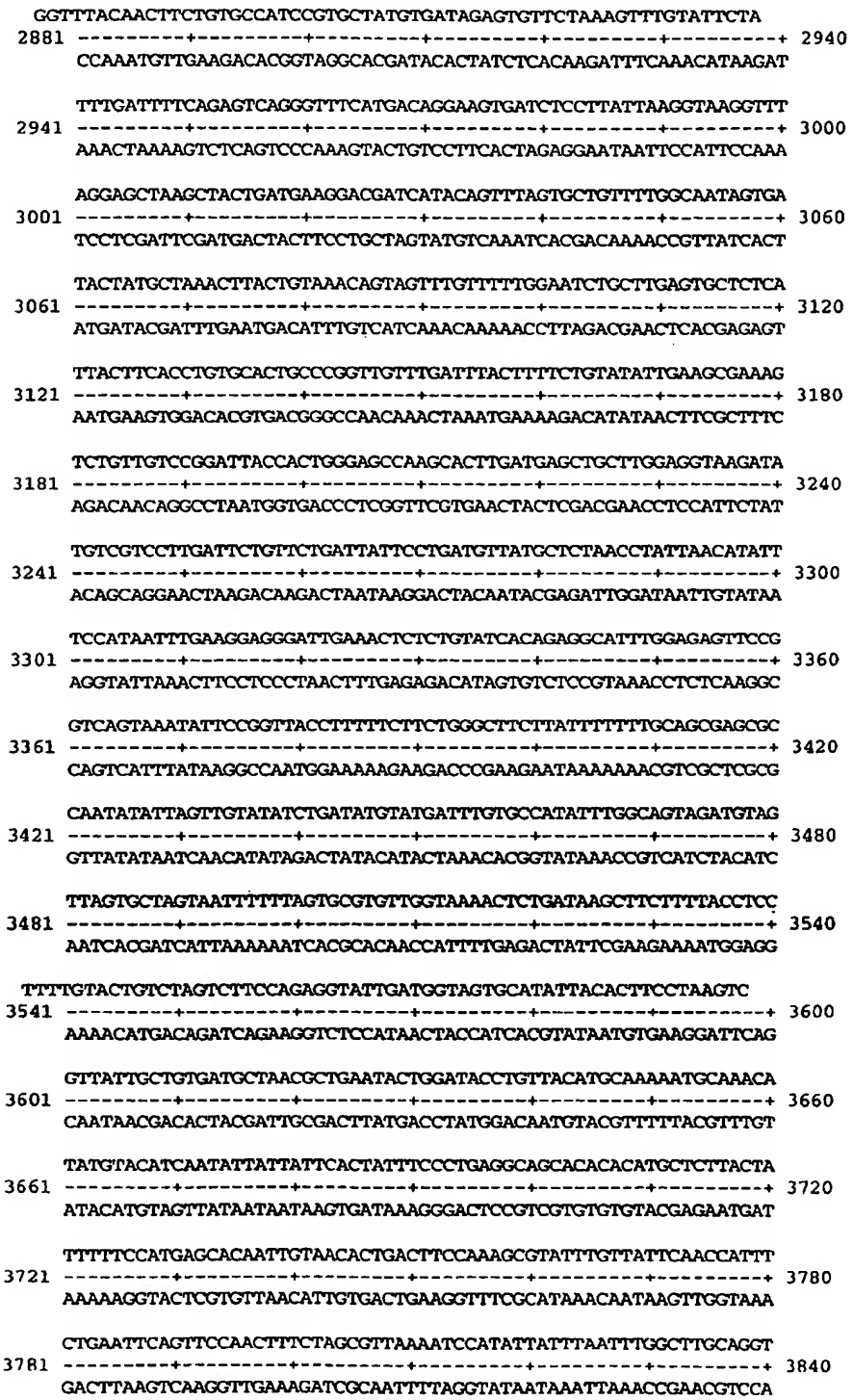


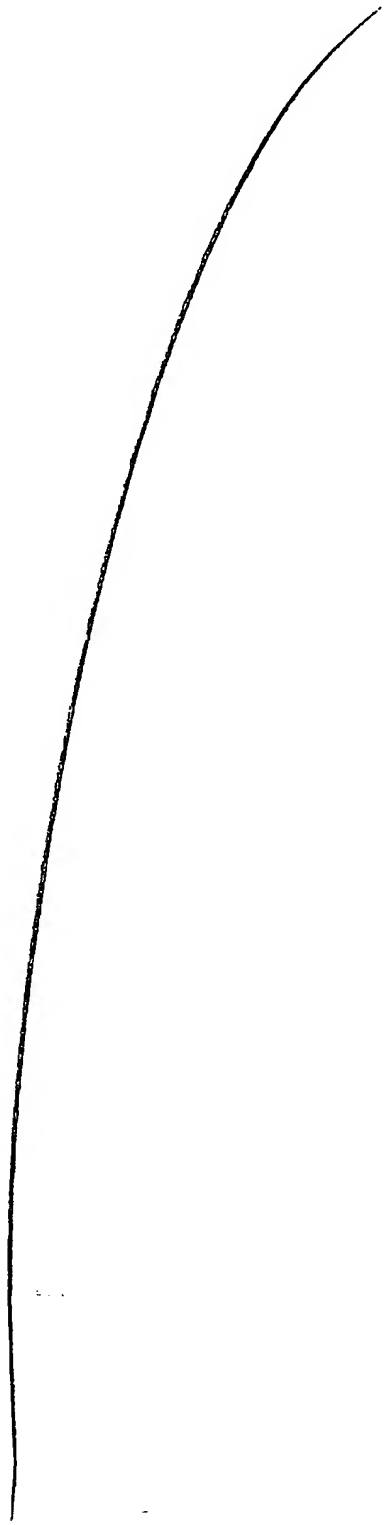
# FIGURE5(A) continued



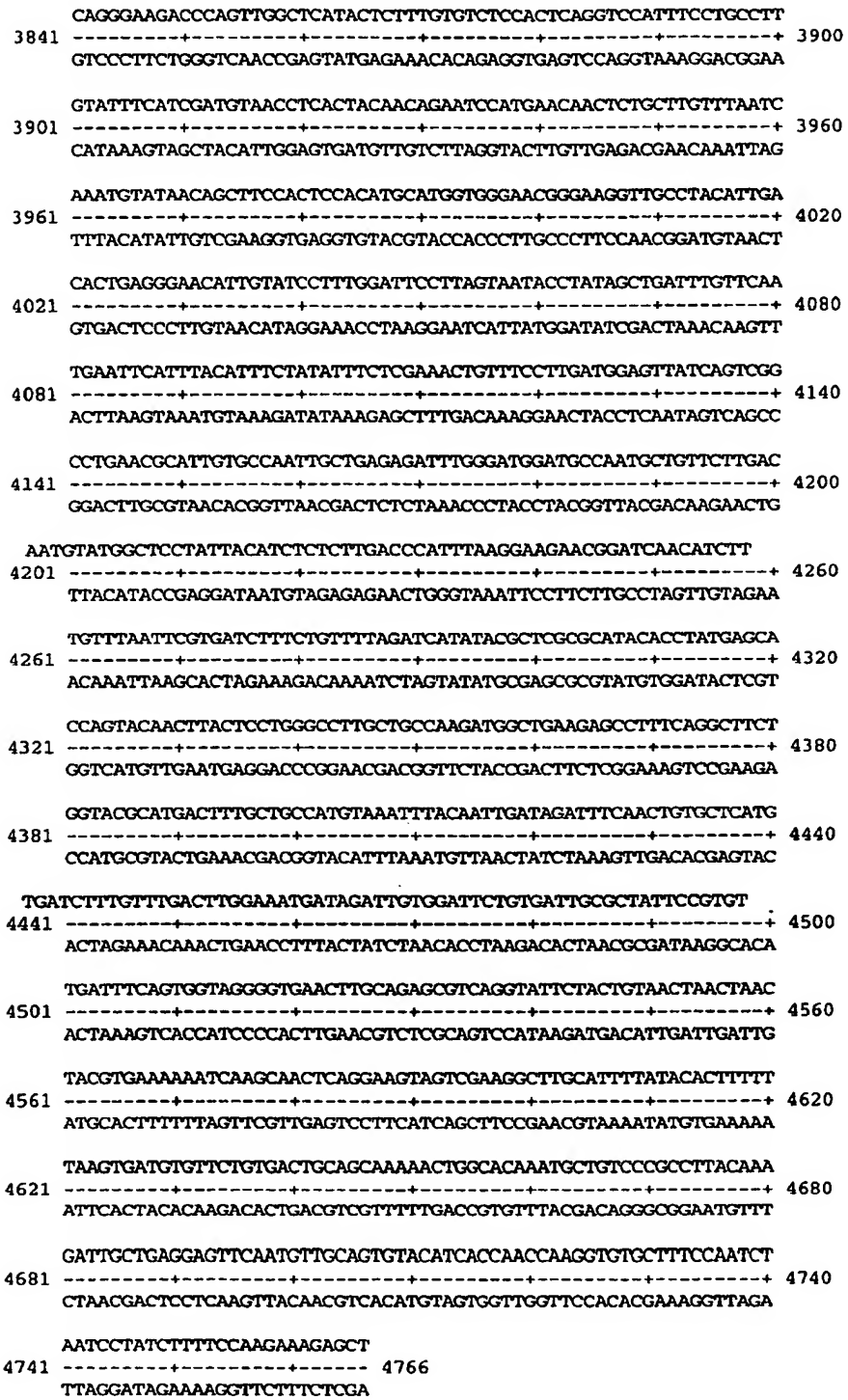


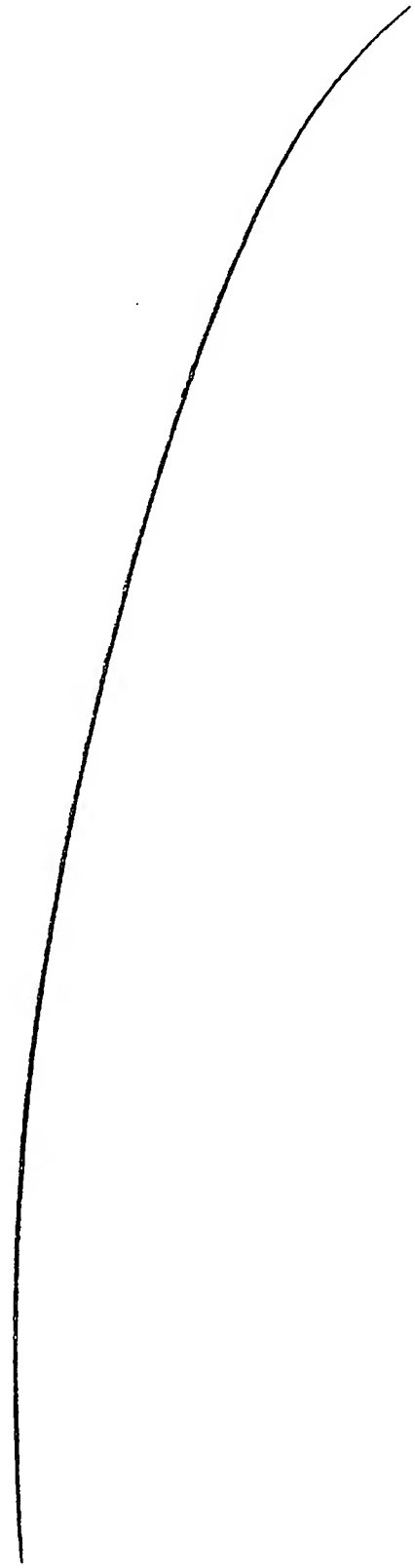
# FIGURE5(A) continued



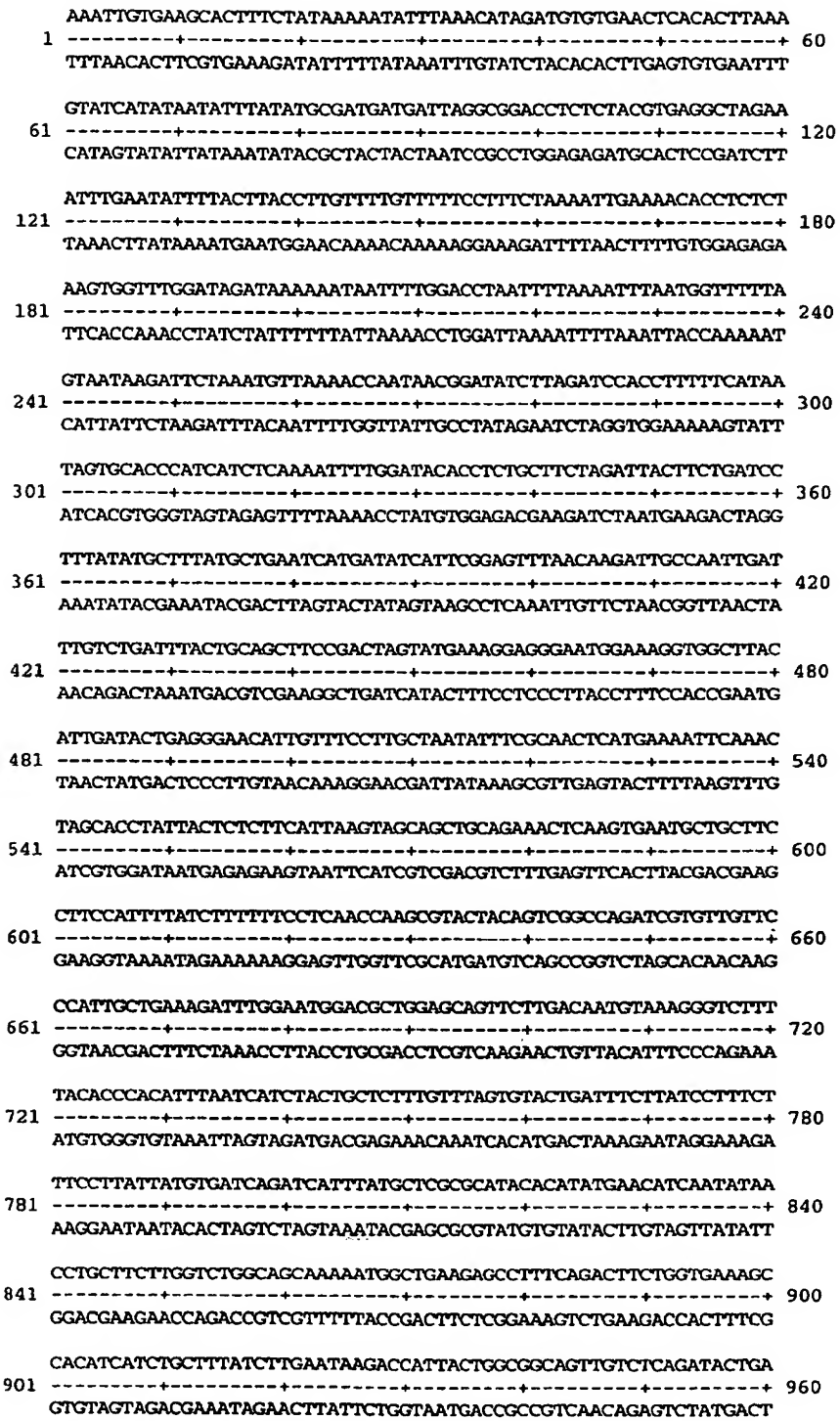


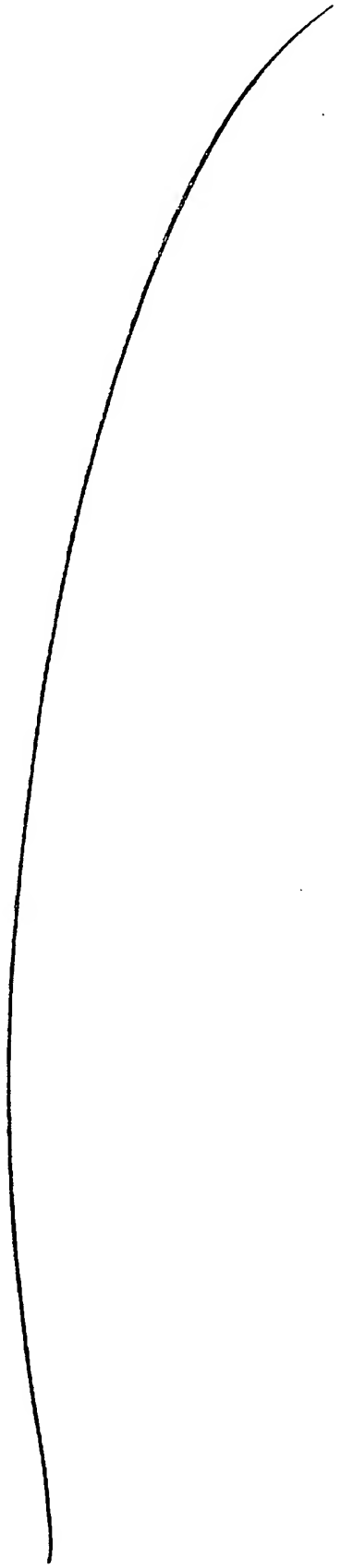
# FIGURE5(A)continued



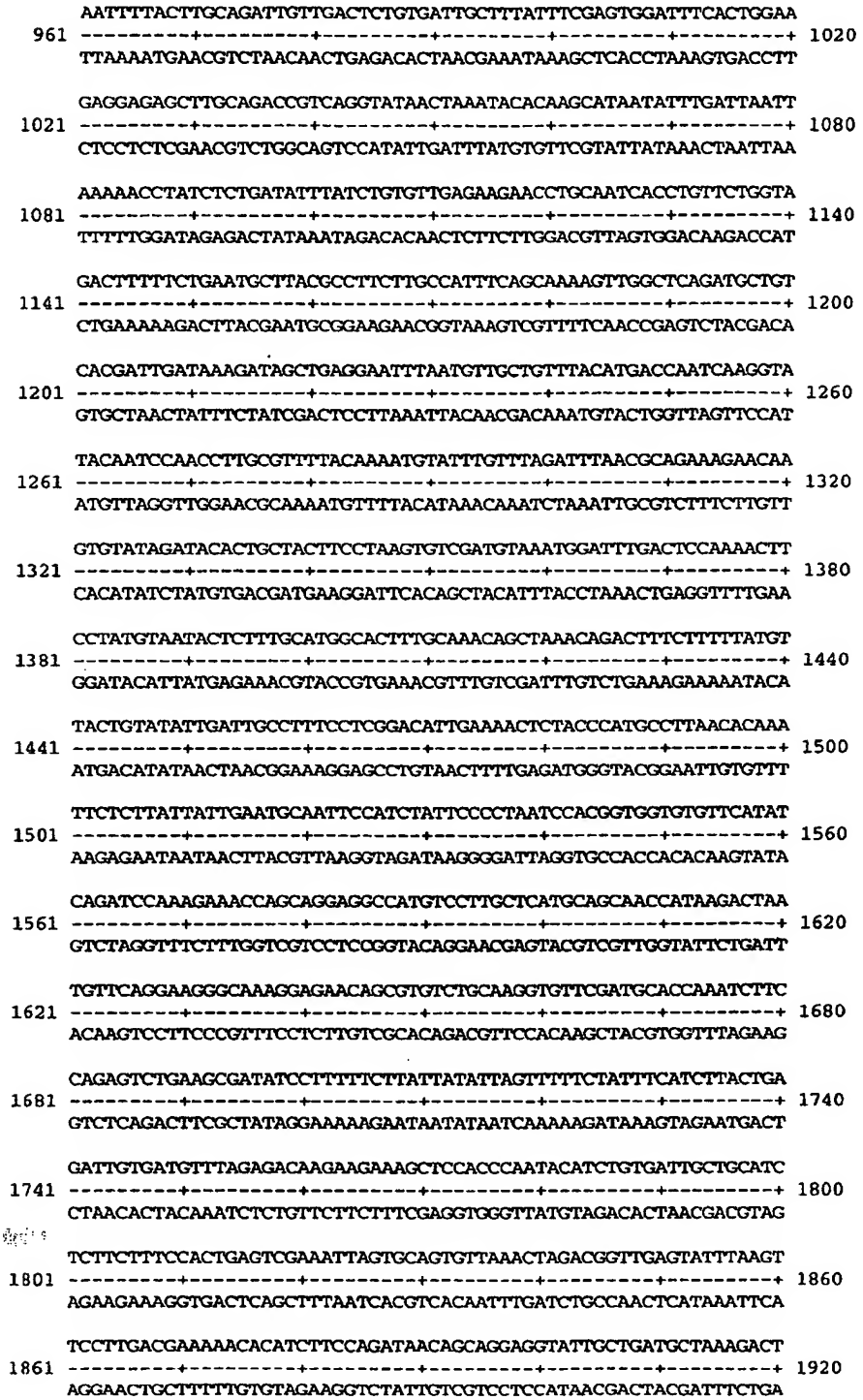


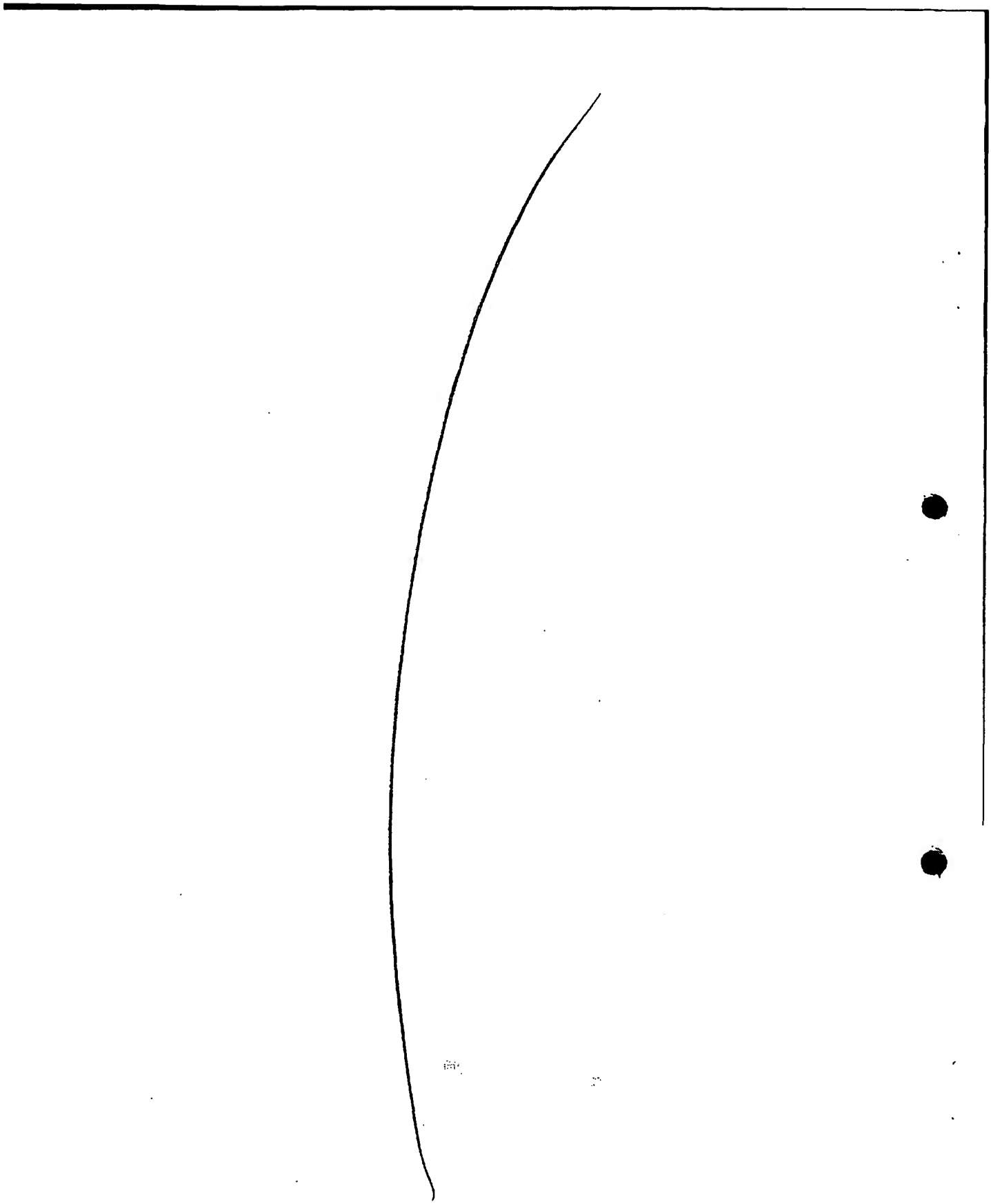
# FIGURE5(B)


AAATTGTGAAGCACTTTCTATAAAAATATTTAAACATAGATGTGTGAACTCACACTTAAA
1 60
TTTAACACTTCGTGAAAGATATTTTTATAAATTTGTATCTACACACTTGAGTGTGAATTT
GTATCATATAATATTTATATGCGATGATGATTAGGCGGACCTCTCTACGTGAGGCTAGAA
61 120
CATAGTATATTATAAATATACGCTACTACTAATCCGCCTGGAGAGATGCACTCCGATCTT
ATTTGAATATTTTACTTACCTTGTTTTTGTTTTTCCTTTCTAAAATTGAAAACACCTCTCT
121 180
TAAACTTATAAAATGAATGGAACAAAACAAAAAGGAAAGATTTTAACTTTTGTGGAGAGA
AAGTGGTTTGGATAGATAAAAAATAATTTTGGACCTAATTTTAAAATTTAATGGTTTTTA
181 240
TTCACCAAACCTATCTATTTTTTATTAAAACCTGGATTAAAATTTTAAATTACCAAAAAT
GTAATAAGATTCTAAATGTTAAAACCAATAACGGATATCTTAGATCCACCTTTTTCATAA
241 300
CATTATTCTAAGATTTACAATTTTGGTTATTGCCTATAGAATCTAGGTGGAAAAAGTATT
TAGTGCACCCATCATCTCAAAATTTTGGATACACCTCTGCTTCTAGATTACTTCTGATCC
301 360
ATCACGTGGGTAGTAGAGTTTTAAAACCTATGTGGAGACGAAGATCTAATGAAGACTAGG
TTTATATGCTTTATGCTGAATCATGATATCATTCGGAGTTTAACAAGATTGCCAATTGAT
361 420
AAATATACGAAATACGACTTAGTACTATAGTAAGCCTCAAATTGTTCTAACGGTTAACTA
TTGTCTGATTTACTGCAGCTTCCGACTAGTATGAAAGGAGGGAATGGAAAGGTGGCTTAC
421 480
AACAGACTAAATGACGTCGAAGGCTGATCATACTTTCCTCCCTTACCTTTCCACCGAATG
ATTGATACTGAGGGAACATTGTTTCCTTGCTAATATTTCGCAACTCATGAAAATTCAAAC
481 540
TAACTATGACTCCCTTGTAACAAAGGAACGATTATAAAGCGTTGAGTACTTTTAAGTTTG
TAGCACCTATTACTCTCTTCATTAAGTAGCAGCTGCAGAAACTCAAGTGAATGCTGCTTC
541 600
ATCGTGGATAATGAGAGAAGTAATTCATCGTCGACGTCTTTGAGTTCACTTACGACGAAG
CTTCCATTTTATCTTTTTTCCTCAACCAAGCGTACTACAGTCGGCCAGATCGTGTTGTTC
601 660
GAAGGTAAAATAGAAAAAAGGAGTTGGTTCGCATGATGTCAGCCGGTCTAGCACAACAAG
CCATTGCTGAAAGATTTGGAATGGACGCTGGAGCAGTTCTTGACAATGTAAAGGGTCTTT
661 720
GGTAACGACTTTCTAAACCTTACCTGCGACCTCGTCAAGAACTGTTACATTTCCCAGAAA
TACACCCACATTTAATCATCTACTGCTCTTTGTTTAGTGTACTGATTTCTTATCCTTTCT
721 780
ATGTGGGTGTAAATTAGTAGATGACGAGAAACAAATCACATGACTAAAGAATAGGAAAGA
TTCCTTATTATGTGATCAGATCATTTATGCTCGCGCATACACATATGAACATCAATATAA
781 840
AAGGAATAATACACTAGTCTAGTAAATACGAGCGCGTATGTGTATACTTGTAGTTATATT
CCTGCTTCTTGGTCTGGCAGCAAAAATGGCTGAAGAGCCTTTCAGACTTCTGGTGAAAGC
841 900
GGACGAAGAACCAGACCGTCGTTTTTACCGACTTCTCGGAAAGTCTGAAGACCACTTTCG
CACATCATCTGCTTTATCTTGAATAAGACCATTACTGGCGGCAGTTGTCTCAGATACTGA
901 960
GTGTAGTAGACGAAATAGAACTTATTCTGGTAATGACCGCCGTCAACAGAGTCTATGACT

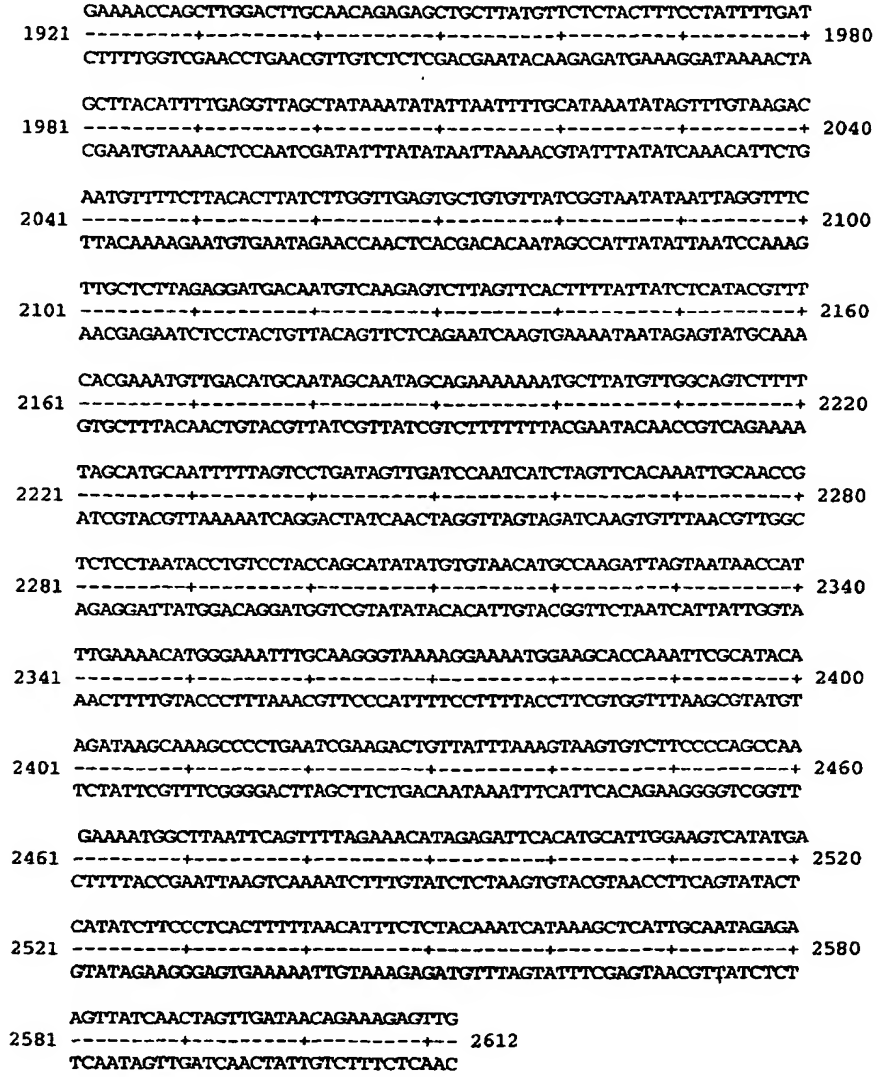


# FIGURE5(B) continued





# FIGURE5(B) continued



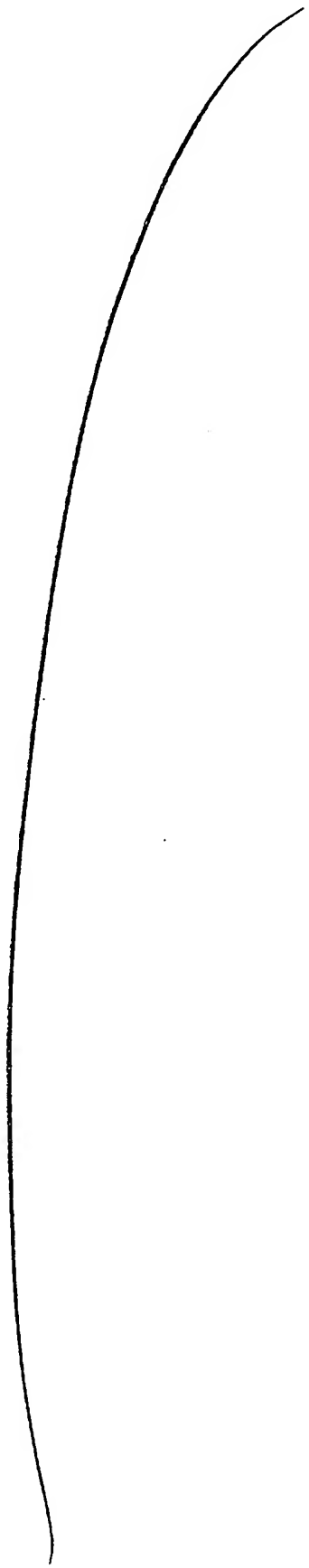
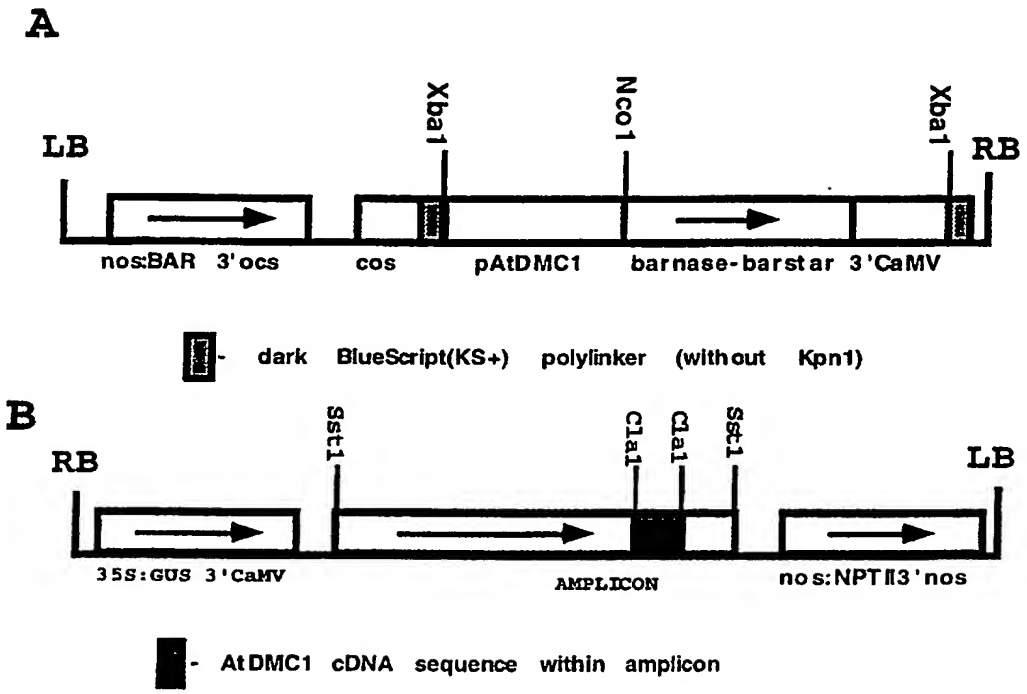
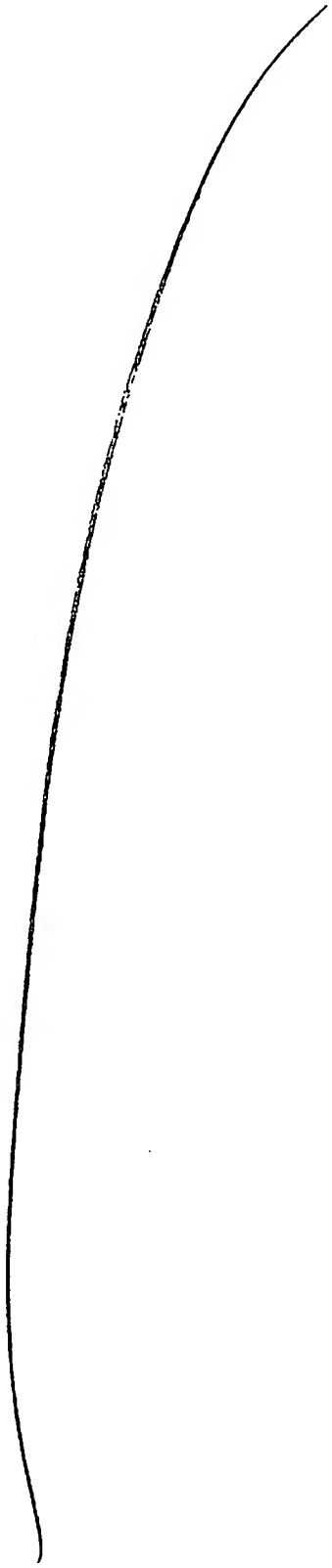


FIGURE 6


A
LB
Xba1
Nco1
Xba1
RB
nos:BAR 3'ocs
cos
pAtDMC1
barnase-barstar 3'CaMV
- dark BlueScript(KS+) polylinker (without Kpn1)
B
RB
Sst1
Cla1
Cla1
Sst1
LB
35S:GUS 3'CaMV
AMPLICON
nos:NPTII3'nos
- AtDMC1 cDNA sequence within amplicon



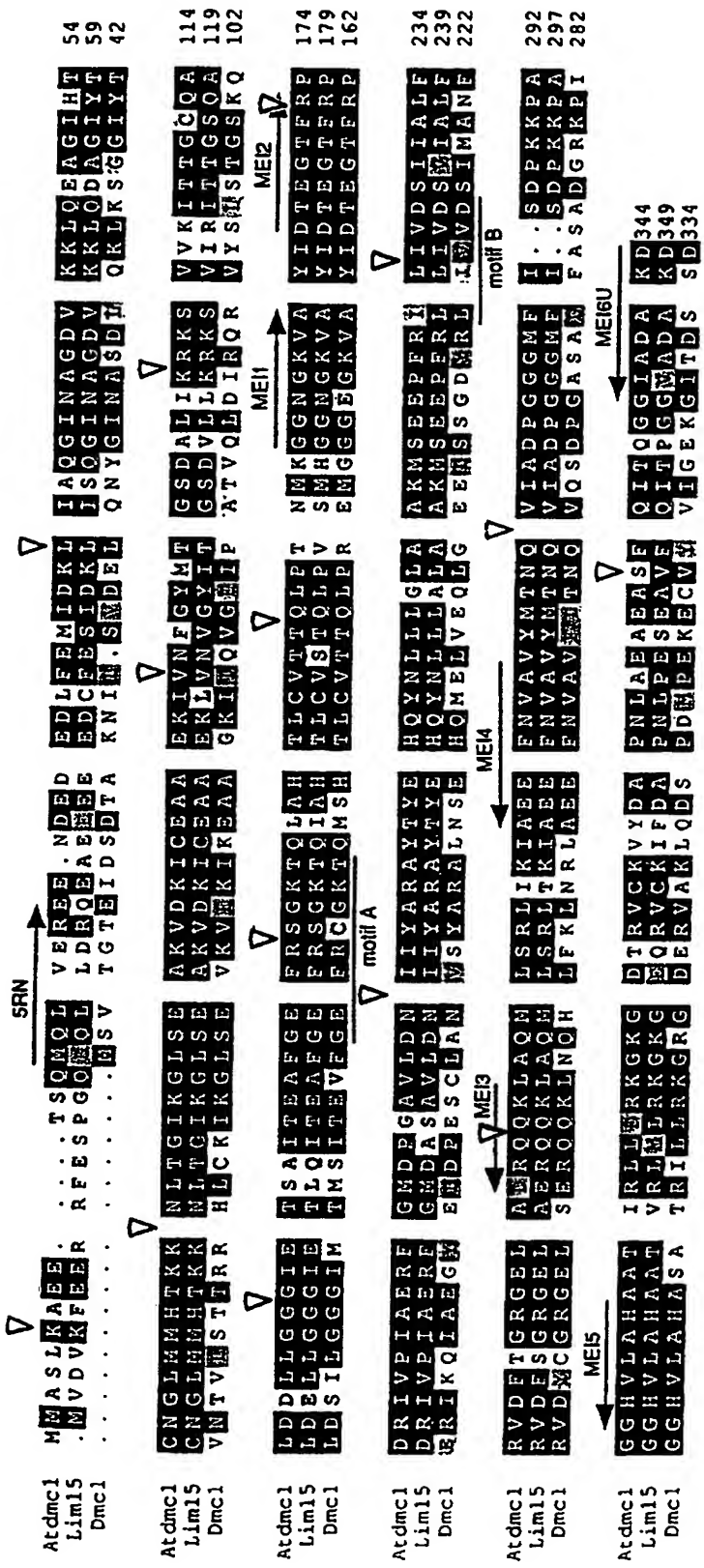
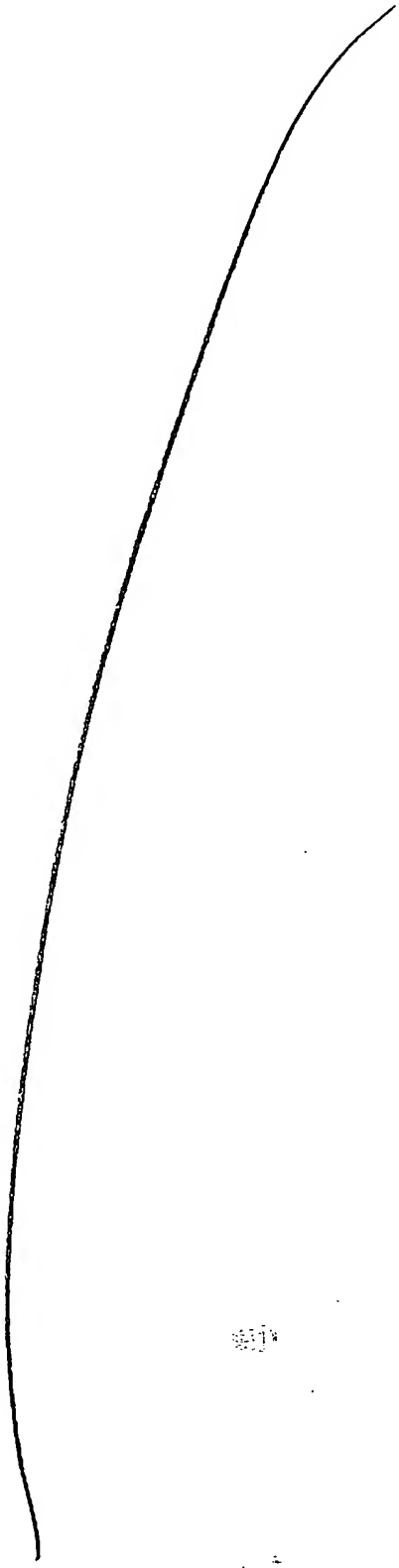

5RN
motif A
MEI3
MEI4
MEI5
MEI1
MEI2
motif B
MEI6U
Atdmc1
Lim15
Dmc1
54
59
42
114
119
102
174
179
162
234
239
222
292
297
282
KD344
KD349
SD334

Figure 7



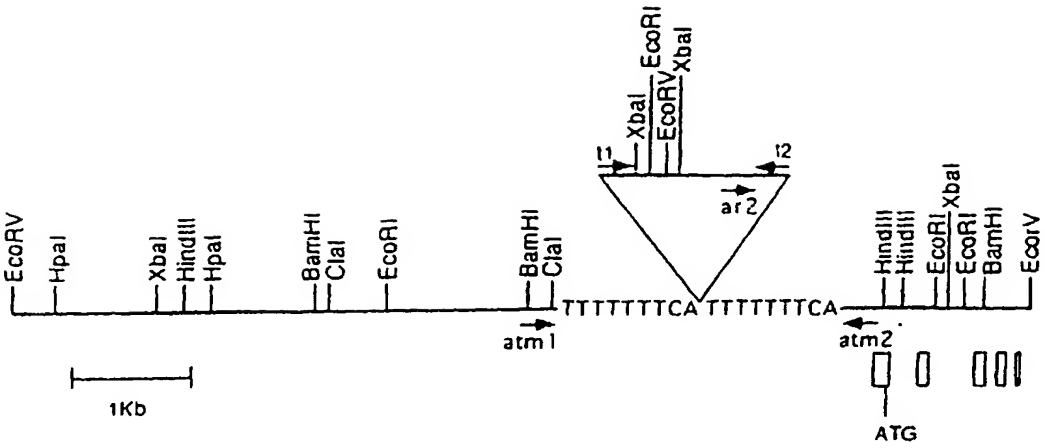

EcoRV
HpaI
XbaI
HindIII
HpaI
BamHI
ClaI
EcoRI
BamHI
ClaI
XbaI
EcoRV
EcoRI
t1
t2
ar2
TTTTTTTCA TTTTTTTCA
atm1
atm2
HindIII
HindIII
EcoRI
XbaI
EcoRI
BamHI
EcoRV
1Kb
ATG

Figure 8